The Dissertation Committee for Michael Drew LaMar
certifies that this is the approved version of the following dissertation:

Human Acoustics: From Vocal Chords to Inner Ear

Committee:

Jack Xin, Supervisor

Irene Gamba

Todd Arbogast

Oscar Gonzalez

Mary Lou Zeeman

Human Acoustics: From Vocal Chords to Inner Ear

by

Michael Drew LaMar, B.S.

Dissertation
Presented to the Faculty of the Graduate School of
the University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin
May 2005

Dedicated to my family.

Acknowledgments

My thanks go out first to my advisor Dr. Jack Xin for his advising, mentoring and friendship. Thanks to Dr. Yingyong Qi for his help and support. I would also like to thank the rest of my thesis committee for their input and advice. I thank my family, especially my Dad for helping support me during my long tenure as a graduate student. Finally, my thanks go out to all of my fellow graduate students.

This research was made possible through a University of Texas Continuing Graduate Fellowship, an ARO Grant No. DAAD 19-00-1-0524 and NSF grant ITR-0219004.

Human Acoustics: From Vocal Chords to Inner Ear

Publication No. ______________

Michael Drew LaMar, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Jack Xin

Part I covers the vocal chords, more accurately known as the vocal folds (VF). Modeling efforts are split into two areas: the VF tissue and the airflow. There are multiple existing models of the VF, with varying ranges of complexity for both the tissue and the airflow. In our model, the tissue is based on a recent two-mass model of Bogaert's [5], while the airflow is quasi-one-dimensional and is derived from the two-dimensional compressible Navier-Stokes equations. Our model is more accurate than Bernoulli's law (quasi-steady approximation), yet less complex than the full Navier-Stokes system. The model is shown to reproduce important transient behaviour intrinsic in vocal fold motion, such as pressure peaks before and after vocal fold closure.

Part II concerns the inner ear, or cochlea. Again the modeling effort is split into two areas: the cochlear tissue and the cochlear fluid. We model the cochlear fluid with the well known two-dimensional box model of the cochlea, derived from the three-dimensional compressible Navier-Stokes equations. The cochlear tissue structure is where the complexity takes place. We start with Neely and Kim's [25] linear active model for the cochlear structure and modify their active gain parameter into a nonlinear nonlocal functional. The nonlinearity forces us to work in the time domain, which is prone to dispersive instabilities if one uses a frequency domain middle ear model. The middle ear's role as a transient absorber is discussed and its time domain formulation is shown to reduce the dispersive instability. We perform simulations on the full system and show that the model recovers many important nonlinear phenomena, such as suppression and difference tones. A spectrogram based on the cochlear response is created and compared with the spectrogram of the input waveform.

In both Part I and Part II, the emphasis is on time dependent modeling and numerical implementation.

Table of Contents

List of Tables

List of Figures

Part I

Vocal Folds

Chapter 1

Introduction

Table 1.1: Part I: List of abbreviations

VF	Vocal folds
NS	Navier-Stokes

The vocal chords, more accurately known as the vocal folds (VF), are the source of the human voice. An accurate mathematical model of the VF could have many applications. In the area of medicine, it could aid in the creation of a prosthetic VF. A model can also be used to possibly diagnose VF abnormalities based on a person's speech waveform (inverse problem). There is also a good deal of communication between singers and VF researchers, showing benefit to the arts.

The intricacies of VF motion are fundamental in speech production and consist of the mechanical interaction between the airflow from the lungs and the elastic response of the VF. In other words, the airflow *drives* the "vibration" of the VFs. We therefore have two structures on which to focus our modeling efforts: the airflow and the VF tissue.

In Chapter 2, we delve into the physiology of voice production, and continue on to the existing experimental work and data to which we will compare our model results. There are some surprising realizations in the biology, one of which is the often misunderstood fundamental motion of the VFs. Voice production, consisting of VF motion and vocal tract resnonance, resembles most the sound production from a brass instrument, not a woodwind instrument.

In Chapter 3, we start by discussing the existing work on the VF and some of the pros and cons of each approach. We then lead into our modeling efforts and show how our model fits into the larger modeling framework. The existing work has fallen into two categories: modeling of the airflow and modeling of the tissue. In each category, there are varying degrees of model complexity, with model complexity generally being proportional to computational effort involved. Our airflow model is a quasi-one-dimensional model, based on the two-dimensional compressible Navier-Stokes equations, that was found using asymptotic methods. It is a new approach in that it simplifies the airflow from the full two-dimensional Navier-Stokes equations, while staying in the time domain by not making the quasi-steady approximation (Bernoulli's Law). This is important in order to capture important transient effects that Bernoulli's Law alone misses. Our tissue model is the two-mass model of Bogaert [5], a recent modification of Ishizaka & Flanagan's two-mass model [17].

In Chapter 4, derivation of numerical methods and issues therein will be discussed. The resulting airflow equations are in the form of a conservation law with source term. We use time-splitting, a first order approximation. Implementation of conservation methods are then discussed, as well as future work in developing higher-order methods.

In Chapter 5, we show results from our simulations and compare them with the data and experiments in Chapter 2. An important achievement of our model is its ability to recover double pressure peaks before and after vocal fold closure. This is a transient phenomenon that is not seen when one uses a quasi-steady flow approximation. In fact, the quasi-steady flow approximation is not accurate during a crucial *one-fifth* of the vocal fold cycle [11, 23], corresponding to opening and closing of the VFs.

Chapter 6 concludes Part I on the vocal folds, summarizing the previous chapters and discussing future work. Much of the work in Part I was first published in [19].

Chapter 2

Vocal Fold Physiology

2.1 Anatomy

A schematic of the physiology of the voice box and surrounding tissue is given in Figure 2.1. The voice box is also called the larynx and consists of a collection of bone, muscle and ligaments. It houses the vocal folds, which are considered the source of voice production. As you can see, the complexity involved in a full mechanical model of the voice box would be a daunting task. The larynx moves in complicated ways in voice production, such as up and down movements in controlling pitch, and has definite effects on the voice output. However, we will concentrate our efforts on the vocal folds, ignoring the movement of the larynx and surrounding bone/muscle structure.

The process of voice production can be broken up into three main steps. The first step takes place in the lungs and is where the airflow is created. The second step is in the glottis (the space between the vocal folds) where the airflow interacts with the VF tissue to create "sound". The third and final stage is the resulting pressure disturbance traveling through the vocal tract, which consists of the oral and nasal cavity, shaping the "sound" through resonant frequencies. An instructive analogy is to think of a musical instrument, for the moment brass or woodwind. In the instrument, there is a mouthpiece (the VF) and the instruments resonant cavity (vocal tract). It is a common misconception to imagine the VF vibrating like that of a woodwind

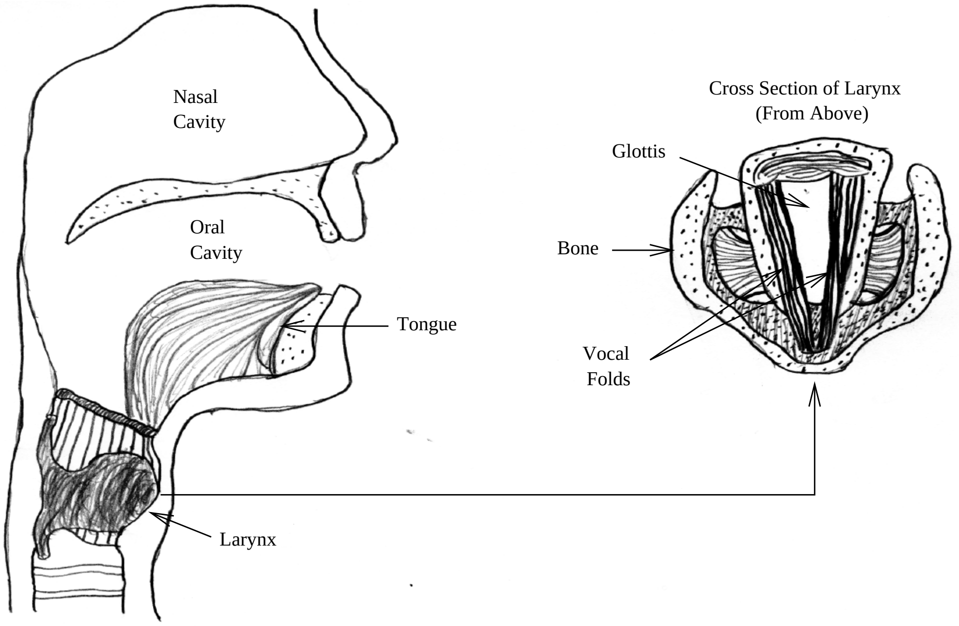

Figure 2.1: A cross-sectional view of the head with a view of the vocal folds within the larynx (Reconstructed from [30]).

mouthpiece. In fact, the VF are more analogous to the vibration of the player's lips in a brass instrument's mouthpiece. The VFs create sound through their opening and closing, which creates puffs of air. They slam together upon closing, staying closed for one-fifth of the cycle. This is why we can actually fatigue our vocal chords. Some singers have gotten around this by training their VFs to not slam together completely during closure.

2.2 Measurements

This section contains measurements of both air pressure and volume velocity (air flux) at different locations in the glottis. Figure 2.2, reproduced from Titze [34], displays experimentally measured intraglottal pressure on an

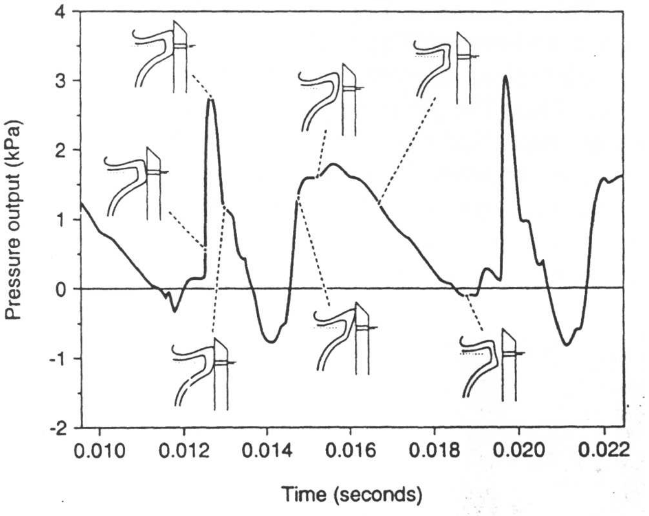

Figure 2.2: Experimentally measured intraglottal pressure on an excised canine larynx (reproduced Fig. 8 on p. 426 of Titze [34] with permission).

excised canine larynx (see also [18]). The double pressure peaks occur during the opening and closing of the vocal folds. Models that make quasi-steady approximations on the airflow are unable to resolve these peaks. We try to recover this qualitative behaviour in our model.

Figure 2.3, reproduced from Story and Titze [33], shows simulated air volume velocity computed at the VF exit using pressure recovery downstream of the separation point. Figure 6 in Alipour and Scherer [1] is a similar plot of air volume velocity computed using a 2D incompressible Navier-Stokes calculation on a numerical domain covering a considerable wake flow region downstream of the vocal fold exit.

Figure 2.4 shows experimenal data collected from human subjects. We will compare our model results with their plot of transglottal pressure, denoted p_{trans}. Note the occurrence of three peaks in pressure in each cycle, two

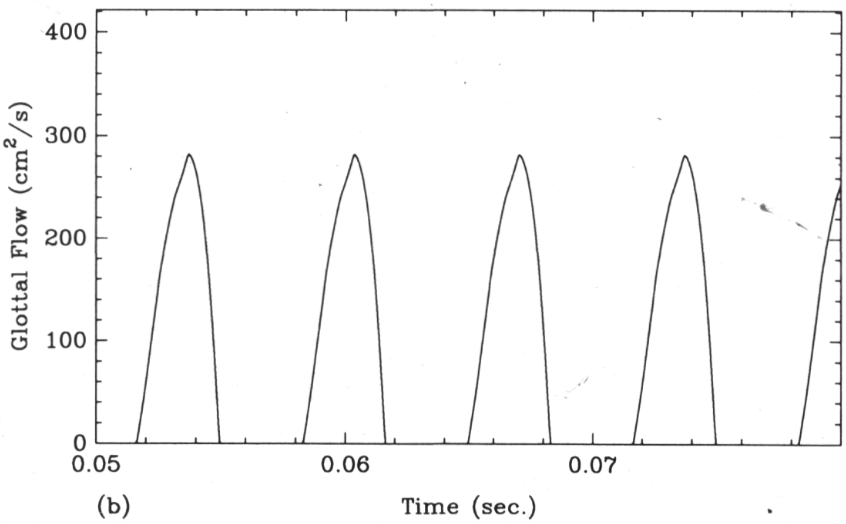

Figure 2.3: Simulated air volume velocity at VF exit using pressure recovery downstream of the separation point (reproduced Fig. 7 from Story and Titze [33] with permission).

corresponding to VF opening (bottom of VF, then top) and the third to VF closure.

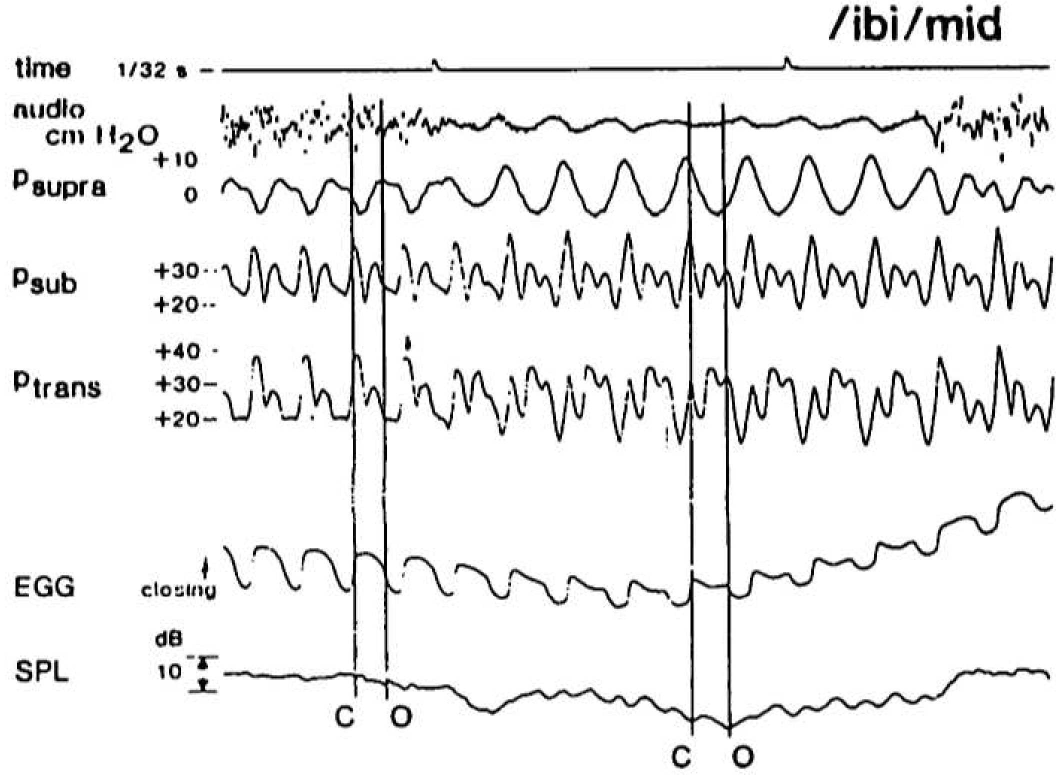

Figure 2.4: Experimental data collected from human subjects [3]. (Reprinted from Journal of Voice (11), Austin and Titze, “The effect of subglottal resonance upon vocal fold vibration”, p. 391–402, Copyright 1997, with permission from The Voice Foundation.)

Chapter 3

Modeling

In this chapter we derive the model equations for the VF motion. We simplify the VF system from three to two dimensions by assuming symmetry along the z-dimension (see Figure 3.1). Then, using the two-dimensional Navier-Stokes equations, we make some assumptions on the flow to derive quasi-one-dimensional flow equations. This is done by assuming the flow is mostly one-dimensional and then averaging the variables along the non-dominant (y) direction. Moving on to the tissue, we discuss a two-mass model of the VF by Bogaert [5], a recent improvement on Ishizaka & Flanagan [17]. Finally, we discuss the way in which the airflow and VF are coupled, as well as modeling intricacies, such as airflow separation and vocal fold closure.

3.1 Airflow: The Modified Euler Equations

We start with three-dimensional flow through a channel given in Figure 3.1, where flow moves from the subglottal region through the VFs to the supraglottal region. Our first simplification will be to impose symmetry along the z-direction and consider two-dimensional flow through the glottis only (see Figure 3.2). The region is defined by

$$\Omega_0 = \Omega_0(t) = \{(x, y) : x \in [-L, L], y \in [-A(x,t)/2, A(x,t)/2]\} \tag{3.1}$$

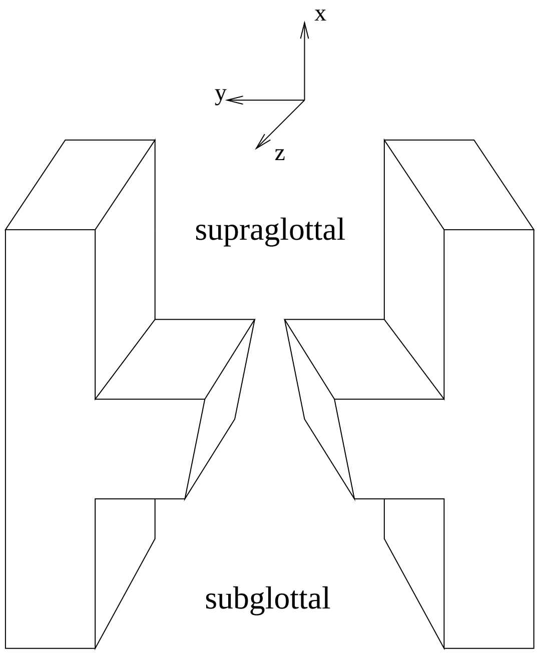

Figure 3.1: Simple picture of the vocal folds and glottis.

where $A = A(x,t)$ denotes glottal width. We will then make assumptions on the flow that will allow us to work with y-averaged quantities of pressure, velocity, and density, thereby reducing the problem to a *quasi*-one-dimensional flow. We then conclude the section with remarks on viscous effects in our model.

We start with two-dimensional flow quantities given by

$$\begin{array}{lllll} Position & = & (x,y) & \equiv & (x_1,x_2) \\ Pressure & = & p & = & p(x,y,t) \\ Density & = & \rho & = & \rho(x,y,t) \\ Velocity & = & \vec{u} & = & (u_1(x,y,t),u_2(x,y,t)) \end{array}$$

The two-dimensional Navier-Stokes equations in differential form are given as

Conservation of mass:

$$\rho_t + \nabla \cdot (\rho \vec{u}) = 0 \tag{3.2}$$

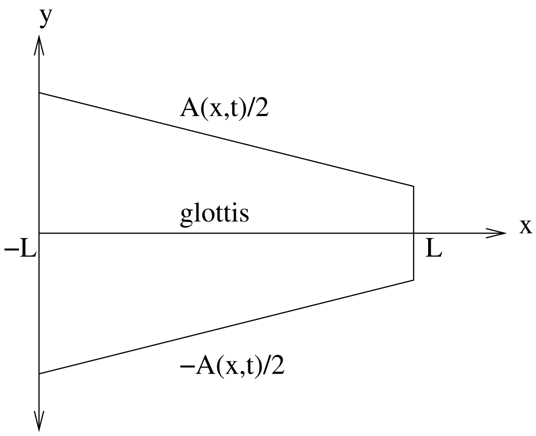

Figure 3.2: Computational domain $\Omega_0(t)$ of the glottis.

Conservation of momentum:

$$(\rho\vec{u})_t = -\nabla \cdot (\rho(\vec{u} \otimes \vec{u})) + \mathrm{div}(\sigma \cdot \vec{n}) \tag{3.3}$$

where σ is the stress tensor given by

$$\begin{aligned} \sigma &= \sigma_{ij} = -p\delta_{ij} + d_{ij}, & (3.4)\\ & \quad\ d_{ij} = 2\mu\left(e_{ij} - \frac{\mathrm{div}\vec{u}}{3}\delta_{ij}\right), & (3.5)\\ & \quad\ e_{ij} = \frac{1}{2}(u_{i,x_j} + u_{j,x_i}) & (3.6) \end{aligned}$$

3.1.1 Derivation

In this section, we will use asymptotic methods to reduce equations (3.2) and (3.3) to quasi-one-dimensional flow equations. The reduced equations will have the following y-averaged quantities:

$$\begin{aligned} \bar{\rho} &= \tfrac{1}{A}\textstyle\int_{-A/2}^{A/2} \rho(x,y,t)\ dy \\ \bar{p} &= \tfrac{1}{A}\textstyle\int_{-A/2}^{A/2} p(x,y,t)\ dy \\ \bar{u}_1 &= \tfrac{1}{A}\textstyle\int_{-A/2}^{A/2} u_1(x,y,t)\ dy \end{aligned}$$

The reduced equations are given by

Conservation of mass:

$$(\bar{\rho}A)_t + (\bar{\rho}\bar{u}_1 A)_x = 0 \tag{3.7}$$

Reduced momentum equation:

$$(\bar{\rho}\bar{u}_1 A)_t + (\bar{\rho}\bar{u}_1^2 A)_x = -(\bar{p}A)_x + A_x\bar{p} + \bar{\rho}\bar{u}_1 A_t \tag{3.8}$$

The fluid will be slightly viscous subsonic air flow with the equation of state either polytropic or isothermal. We define a volume element $\Omega(t) \subset \Omega_0(t)$ of the form

$$\Omega(t) = \{(x, y) : x \in [a, b] \subset [-L, L], y \in [-A(x, t)/2, A(x, t)/2]\}$$

The boundary conditions for the upper and lower boundary stem from the velocity no slip boundary condition and are given by

$$y = \pm A(x, t)/2 : \quad \rho_y = 0,\ \vec{u} = (0, \pm A_t/2) \tag{3.9}$$

At the inlet $x = -L$, we set the pressure to $p = p_0$ and flow velocity to $\vec{u} = (u_{1,0}, u_{2,0})$. At the VF exit $x = L$, we impose a Neumann condition $(p, u_1, u_2)_x = 0$ to help the waves move freely out of the domain.

We will do vertical averaging of the flow to reduce the complexity of the system. This will only be valid under certain assumptions on the flow, which will be presented next. First, for small viscosity (in the end, we take the inviscid limit), the flows are laminar in the interior of Ω_0, forming viscous boundary layers near the VFs. The vertical averaging of the flow will be much less influenced by this boundary layer behaviour as long as $A(x, t) \gg$

$O(\mu^{1/2})$. We assume the flow gradient is mainly one-dimensional away from the boundary layers of Ω_0, i.e.

$$\begin{aligned} |u_{1,y}| &\ll |u_{1,x}|, \\ |u_{2,y}| &\ll |u_{1,x}|. \end{aligned}$$

In the viscous boundary layers (of width $O(\mu^{1/2})$), we have

$$|\vec{u}_y| \gg |\vec{u}_x|$$

Throughout Ω_0,

$$|\rho_y| \ll |\rho_x|$$

These assumptions are consistent with observations in viscous boundary layers of large vertical velocity gradients with small density/pressure gradients [37].

Our intention is to take vertical averages, so denote $\bar{\rho}$, $\bar{u_1}$ as the vertical averages of density and x-component of flow velocity. For calculation purposes, note that the exterior normal $\vec{n}$ to the VF is given by

$$y = \pm A/2 : \quad \vec{n} = (-A_x/2, \pm 1)/\sqrt{1 + A_x^2/4} \tag{3.10}$$

This completes the assumptions and setup.

We will perform our analysis on an infinitesimally small volume element $\Omega(t)$ with $a = x$, $b = x + \delta x$, $t = t + \delta t$ and $\delta x \ll 1$, $\delta t \ll 1$. We will first derive the modified conservation of mass equation (3.7). Letting $J(t)$ denote the Jacobian of area change from $t_0 \to t$, we have

$$\begin{aligned} \frac{d}{dt}\int_{\Omega(t)} \rho \, dV &= \frac{d}{dt}\int_{\Omega(t_0)} \rho J(t) \, dV \\ &= \int_{\Omega(t_0)} \rho_t J \, dV + \int_{\Omega(t_0)} \rho J_t \, dV \end{aligned} \tag{3.11}$$

Since $\delta x \ll 1$, $J(t) \approx A(t)/A(t_0)$, and thus the second integral on the right hand side in (3.11) becomes

$$\int_{\Omega(t_0)} \rho J_t \; dV = \bar{\rho}\frac{A_t(t)}{A(t_0)}A(t_0)\delta x = \bar{\rho}A_t\delta x \tag{3.12}$$

For the first integral in (3.11), we substitute the conservation of mass equation and integrate by parts, arriving at

$$\int_{\Omega(t_0)} \rho_t J(t) \; dV = \int_{\Omega(t)} \rho_t \; dV = -\int_{\partial\Omega(t)} \rho\vec{u}\cdot\vec{n} \; ds \tag{3.13}$$

Using the definition of the normal $\vec{n}$ in (3.10) and arc length $ds = \sqrt{1 + A_x^2/4}\, dx$, we arrive at

$$\begin{aligned}
\int_{\partial\Omega} \rho\vec{u}\cdot\vec{n} \; ds &= \int_{-A/2}^{A/2} (\rho u_1)(s,y,t) \; dy\Bigg|_{s=x}^{s=x+\delta x} \\
&\quad + \int_x^{x+\delta x} \rho\cdot(0, A_t/2)\cdot(-A_x/2, 1) \; dx \\
&\quad + \int_x^{x+\delta x} \rho\cdot(0, -A_t/2)\cdot(-A_x/2, -1) \; dx \\
&= \overline{\rho u_1}A|_x^{x+\delta x} + \frac{\delta x}{2}(\rho A_t)|_{y=A/2} + \frac{\delta x}{2}(\rho A_t)|_{y=-A/2} + O(\delta x^2) \\
&\approx (\bar{\rho}\cdot\bar{u_1}A)|_x^{x+\delta x} + \bar{\rho}A_t\delta x + O(\delta x^2) \qquad (3.14)
\end{aligned}$$

At the boundary $A = \pm A/2$, $\rho_y = 0$ and thus $\rho|_{y=\pm A/2}$ can be approximated by $\bar{\rho}$. Also, $|\rho_y| \ll |\rho_x|$, and thus we can approximate $\overline{\rho u_1}$ by $\bar{\rho}\bar{u}_1$. We also have

$$\int_\Omega \rho \; dV = \bar{\rho}A\delta x + O(\delta x^2) \Longrightarrow \frac{d}{dt}\int_\Omega \rho \; dV = (\bar{\rho}A\delta x)_t + O(\delta x^2) \tag{3.15}$$

Finally, combining (3.11)–(3.15) and picking out the first-order terms in δx, we have

$$(\bar{\rho}A)_t + (\bar{\rho}\bar{u}_1A)_x = 0$$

thus arriving at equation (3.7).

Looking now at the momentum equation (3.3), let us fix $i = 1$, $a = x$ and $b = x + \delta x$. Similar to (3.11), we have

$$\frac{d}{dt}\int_{\Omega(t)} \rho u_1 \; dV = \int_{\Omega(t)} (\rho u_1)_t \; dV + \int_{\Omega(t_0)} \rho u_1 J_t \; dV$$

The second integral is done similar to (3.12) to arrive at

$$\int_{\Omega(t_0)} \rho u_1 J_t \; dV \approx \bar{\rho}\bar{u_1} A_t \delta x + O(\delta x^2)$$

The first integral, similar to (3.13), becomes

$$\int_{\Omega(t)} (\rho u_1)_t \; dV = -\int_{\partial\Omega(t)} \rho u_1 \vec{u} \cdot \vec{n} \; ds + \int_{\partial\Omega(t)} \sigma_{1,j} \cdot \vec{n_j} \; ds$$

Similar to (3.15), we also have

$$\frac{d}{dt}\int_{\Omega} \rho u_1 \; dV = (\overline{\rho u_1} A)_t \delta x + O(\delta x^2) \approx (\bar{\rho} \cdot \bar{u_1} A)_t \delta x + O(\delta x^2)$$

Similar to the calculation in (3.14), using the fact that $u_1 = 0$ on the upper and lower boundaries, we have

$$\int_{\partial\Omega} \rho u_1 \vec{u} \cdot \vec{n} \; ds = (\bar{\rho}\bar{u}_1^2 A)|_x^{x+\delta x} + O(\delta x \mu^{1/2}) \tag{3.16}$$

The smallness of $u_{1,y}$ relative to $u_{1,x}$ in the interior and the small width of the boundary layer $O(\mu^{1/2})$ gives $O(\delta x \mu^{1/2})$ in approximating $\bar{u_1^2}$ by $\bar{u_1} \cdot \bar{u_1}$. Continuing our calculation, we have

$$\begin{aligned} \int_{\partial\Omega} -p\delta_{1,j} n_j \; ds &\approx -\bar{p}A|_x^{x+\delta x} + \int_x^{x+\delta x} pA_x \; dx \\ &= -\bar{p}A|_x^{x+\delta x} + \bar{p}A_x \delta x + O((\delta x)^2). \end{aligned}$$

By definition

$$d_{11} = 2\mu(u_{1,x} - (u_{1,x} + u_{2,y})/3), \quad d_{12} = \mu(u_{1,y} + u_{2,x})$$

By (3.9), we have

$$\bar{u}_{2,y} = \frac{1}{A}\int_{-A/2}^{A/2} u_{2,y}\, dy = \frac{A_t}{A}$$

giving

$$\overline{d_{11}} = \frac{4}{3}\mu\bar{u}_{1,x} - \frac{2\mu A_t}{3A}$$

Thus the contribution from the left and right boundaries located at x and $x+\delta x$ is

$$\sum_{l,r}\int_{l,r} d_{11}n_1 = A\overline{d_{11}}|_x^{x+\delta x} = \frac{4}{3}A\mu\bar{u}_{1,x}|_x^{x+\delta x} - \left.\frac{2\mu A_t}{3}\right|_x^{x+\delta x}$$

The upper and lower boundary contributions give

$$\begin{aligned}\sum_{\pm}\int_{y=\pm A/2} d_{11}n_1\, ds &= -d_{11}A_x\delta x/2|_{y=A/2} - d_{11}A_x\delta x/2|_{y=-A/2} \\ &= \mu\delta x\sum_{\pm} O(\partial_y\vec{u})|_{y=\pm A/2} \qquad (3.17)\end{aligned}$$

We also have,

$$\sum_{\pm}\int_{y=\pm A/2} d_{12}n_2\, ds = \mu\delta x\sum_{\pm} O(\partial_y\vec{u})|_{y=\pm A/2} \qquad (3.18)$$

Since $\partial_y\vec{u}|_{y=\pm A/2} = O(\mu^{-1/2})$, the viscous flux from the boundary layers are $O(\delta x\mu^{1/2})$, much larger than the averaged viscous term $\delta x(4\mu/3)(A\overline{u_{1,x}})_x = O(\delta x\mu)$. Notice that the vertically averaged quantities have little dependence on the viscous boundary layers unless A is on the order $O(\mu^{1/2})$. Hence the quantities from upper and lower edges in (3.17) and (3.18), along with (3.16), should balance themselves. Omitting them altogether, and combining remaining terms that involve only $\overline{u_1}$, $\bar{\rho}$ in the bulk, we end up with (after dividing by δx and sending it to zero)

$$(\bar{\rho}\bar{u}_1A)_t + (\bar{\rho}\bar{u}_1^2A)_x = -(\bar{p}A)_x + A_x\bar{p} + \bar{p}\bar{u}_1A_t + \frac{4\mu}{3}(A\overline{u_{1,x}})_x - 2\mu A_{tx}/3 \quad (3.19)$$

which gives the modified Euler (3.8) in the inviscid limit $\mu \to 0$.

For the remaining part of this paper, we will denote the y-averaged quantities as $\rho \equiv \bar{\rho}$, $p \equiv \bar{p}$ and $u \equiv \bar{u}_1$. Thus, we will denote equations (3.7) and (3.8) as

$$(\rho A)_t + (\rho u A)_x = 0 \tag{3.20}$$

$$(\rho u A)_t + (\rho u^2 A)_x = -(pA)_x + A_x p + \rho u A_t \tag{3.21}$$

3.1.2 Viscous Effects

The viscous effect in the flow produced the term $\rho u A_t$ from the no-slip boundary condition (3.9). Eliminating this term gives the quasi-one-dimensional Euler equations from gas dynamics. This term is very important in transferring energy to the VF from the airflow. There are other viscous effects that we ignored for model simplicity.

3.2 Tissue: Two-Mass Model

In the previous section, we simplified the airflow from the full two-dimensional compressible Navier Stokes equations (3.2)–(3.6) to a quasi-one dimensional equations (3.20),(3.21) involving y-averaged pressure and velocity. The geometry of the vocal folds are coupled in these equations through the term $A(x, t)$. We will simplify the geometry of the VF from fully continuous to linear in x, using the relatively recent two-mass model of Bogaert [5]. See Figure 3.3. At the entrance and exit of the glottis, there are two masses, m_1 and m_2, that are constrained to move in the vertical direction only. They are coupled together through a spring k_{12}, with each anchored to a wall through a spring and damper k_i, c_i respectively. The air pressure will be integrated throughout the glottis and enforced solely on m_1, with the motion of m_2 determined strictly through its coupling to m_1. The equations are given by

$$\begin{aligned} m_1 y_1'' + c_1 y_1' + k_1(y_1 - y_{0,1}) + k_{12}(y_1 - y_2 + y_{0,12}) = F \\ m_2 y_2'' + c_2 y_2' + k_2(y_2 - y_{0,2}) + k_{12}(y_2 - y_1 - y_{0,12}) = 0 \end{aligned} \tag{3.22}$$

where

$$F = L_g \int_{x_0}^{x_s} p \; dx \tag{3.23}$$

y_i denotes distance of m_i from the center of the glottis, $y_{0,i}$ denotes the resting distance of m_i so there is no tension on k_i, and $y_{0,12}$ denotes the resting distance for spring k_{12}. It should be noted that Figure 3.3 is a little misleading in that equation (3.22) shows the static deflection force for k_{12} depending strictly on the y-coordinates only, whereas the figure shows the force depending on both x and y-coordinates. Also note that we assume symmetry along the y-direction and all y-coordinates are considered zero at the center of the glottis. L_g is the vocal fold length in the z-direction (see Figure 3.1) and x_s is called the flow separation point, which will be discussed in the next section.

3.2.1 Flow Separation

In the original VF model by Ishizaka & Flanagan [17], the flow separation was always at the VF exit, or x_2. The two-mass model above is a recent improvement in that the flow separation point depends on the glottal geometry, as will be shown below. Flow separation refers to a change in flow behaviour from attachment to the VF wall via a viscous boundary layer to a free jet with vortices and turbulence in its wake. The vortices cause pressure near the wall to be low, and thus can be approximated by setting it to zero or ambient pressure. The two main geometries are a converging glottis ($y_1 > y_2$) and a diverging glottis ($y_1 < y_2$). It is easily seen that flow separation will

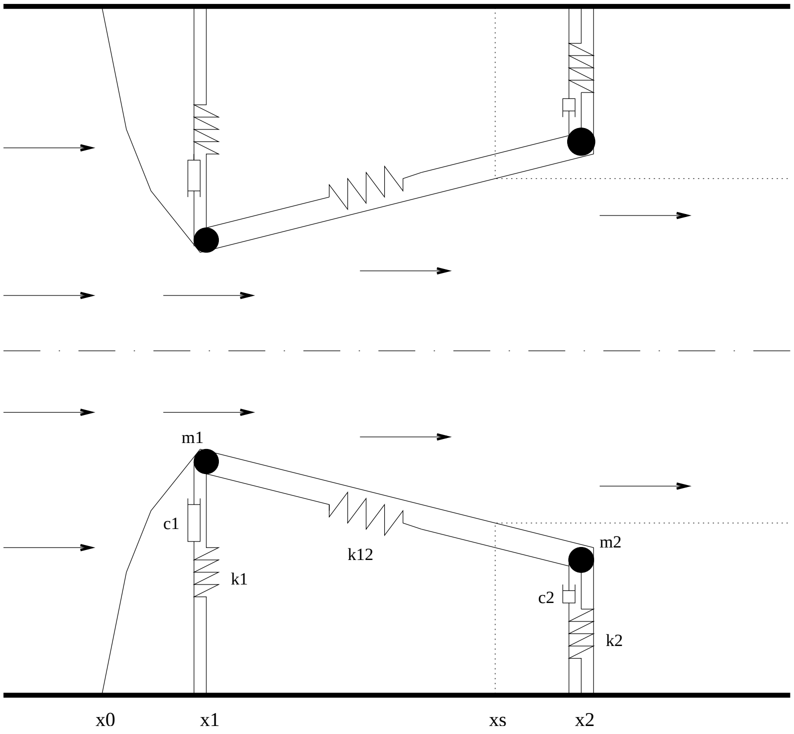

Figure 3.3: A sketch of the airflow and two-mass model. The channel width is fixed at x_0 and piecwise-linearly interpolated from x_0 to x_1. x_s is the separation point of the airflow.

occur at x_2 in a converging glottis, but if the VF are divergent enough (past a certain angle), then flow separation occurs. We use the following empirical formula based on experiments [5, 26]

$$\begin{array}{l} y_2/y_1 < 1.1 \Rightarrow x_s = x_2 \\ y_2/y_1 > 1.1 \Rightarrow x_s = x_1 + \frac{(x_2-x_1)y_1}{10(y_2-y_1)}, \quad y_s = 1.1y_1 \end{array} \tag{3.24}$$

So we see that the separation point depends on the diverging angle. Pressure after the separation point does not contribute force to m_2, so model errors from ignoring viscous losses in the boundary layers are minimized within our system. Also note that viscous effects can be neglected upstream of the separation point for vocal flows [26].

The separation point criteria does not use any information on the reduced flow equations, in particular the viscous boundary layer effects on the flow. A more accurate treatment, therefore, would be to formulate a criterion for the separation point based on contributions from the viscous boundary layer. This type of treatment is performed in [26]. For simplicity, however, we will use equation (3.24).

3.2.2 Vocal Fold Closure

We implement the elastic collision criteria for VF closure [5, 17]. This is done by defining a critical level y_c where the VF are considered closed. When the VF are open, the damping and stiffness constants are set to open parameters. When the VF close, we change the damping and stiffness parameters for each mass to closed parameters (more damping, more stiffness). Thus, we have two VF states:

VF Open: $y_i > y_c \Longrightarrow F = L_g \int_{x_0}^{x_s} p \; dx, \quad k_i = k_{i,open}, \quad c_i = c_{i,open}$
VF Closed: $y_i \leq y_c \Longrightarrow F = L_g \int_{x_0}^{x_1} p \; dx, \quad k_i = k_{i,closed}, \quad c_i = c_{i,closed}$

When the folds are closed, we solve the flow equations over $[x_0, x_1]$ only, and the pressure force on the lower mass becomes

$$L_g \int_{x_0}^{x_1} p \; dx$$

Constant input pressure at the inlet causes pressure to build up and eventually reopen the folds. Note that the two-mass ODE equations are still running during VF closure.

3.3 Model Differences

So what is the major difference between this model and that of Bogaert [5]? The difference is in the airflow where we don't make quasisteady approximations. Instead, we integrate a time-dependent system, which is capable of capturing the transient phenomenon near vocal fold opening and closure that the quasisteady approximation is unable to resolve. To see this relationship a little more clearly, let us scale the velocity to make it nondimensional. Thus, let $v = u/a$, where a is the speed of sound. Substituting into (3.20) and (3.21), and using the equation of state $p = a^2\rho$, we arrive at

$$\begin{array}{l} \frac{1}{a}(Ap)_t + (pvA)_x = 0, \\ \frac{1}{a}(pvA)_t + (pv^2A)_x = -(pA)_x + A_xp + pvA_t/a \end{array} \tag{3.25}$$

We have subsonic flow, and thus $v = u/a \approx 0.1$, the well known Mach number. Using cm · g · ms units, $a = 35$ cm/ms, and thus $1/a$ is a "small" number. If we approximate the equations by neglecting the terms with $1/a$, we have exactly Bernoulli's law for quasisteady flows. As was stated earlier, however, these "small" terms are essential during VF opening and closing.

Chapter 4

Numerics

The flow system (3.25) is of the form

$$U_t + (F(U))_x = G(p, v, A)$$

where $U = (u_1, u_2) = (\frac{1}{a}Ap, \frac{1}{a}pvA)$. This is a conservation law with flux function

$$F(U) = \begin{pmatrix} au_2 \\ au_1 + a\frac{u_2^2}{u_1} \end{pmatrix}$$

and lower order source term

$$G(p, v, A) = \begin{pmatrix} 0 \\ A_x p + pvA_t/a \end{pmatrix}$$

4.1 Time-Splitting

In order to numerically treat the above equation, we use time-splitting, a first-order approximation. We solve the conservation law

$$U_t + (F(U))_x = 0 \tag{4.1}$$

over a full time step. We then update the solution by solving the source term

$$U_t = G(p, v, A) \tag{4.2}$$

and two mass equations (3.22) over the same step. We use an implicit finite difference scheme to numerically solve the source term (4.2) and two-mass equations (3.22). This is important since the ODE's become stiff near vocal fold closure.

The use of time-splitting allows us to use the many results and numerical methods available for conservation laws (see [21]). We use explicit Lax-Friedrichs for our simulations to solve (4.1), which is given by

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{k}{2h}(F(U_{j+1}^n - F(U_{j-1}^n)) \tag{4.3}$$

where k and h are time step and spatial grid size, respectively. The time step k must be small enough to ensure stability of the difference scheme and to keep the computed flow velocity positive, i.e. no back flow is allowed. We use an adaptive time step to ensure this is accomplished in an efficient manner.

4.2 Gudonov's Method and Higher Order Methods

In future work, we would like to implement the conservation law (4.1) numerically using Godunov's method. Godunov's method has less numerical diffusion and spreading effects then Lax-Friedrichs, and is also a stepping stone for many higher-order methods. One such method is the quasi-steady wave propagation algorithm by LeVeque [22], where time-splitting is unnecessary.

Chapter 5

Results

In this section we display our simulation results, comparing them to experimental measurements and other model calculations. Let us first look at the VF motion. Figure 5.1 displays a simulation of a vocal fold cycle and is similar in nature to the figure on p. 113 of [30].

Figure 5.2 displays the displacements y_1, y_2 of the VF masses during 3 cycles of VF vibration. Notice that since all force is placed on m_1, the movement of m_2 is only through the coupling spring k_{12} and therefore the plot of m_2 is nearly a translation of m_1.

We compute the air volume velocity at the exit of the VF and display the results in Figure 5.3. This is similar to the numerical results in Figure 6, Figure 7 and Figure 8 of Story and Titze [33]. Figure 7 is reproduced in our Figure 2.3. It is also similar to Figure 6 of Alipour and Scherer [1]. All of these results come from a more complicated treatment of flow separation. In particular, Story and Titze used pressure recovery downstream of the separation point [33], while Alipour and Scherer used a 2D incompressible NS calculation on a numerical domain covering a considerable wake flow region beyone x_2 [1]. What this shows is that our simplifications recover qualitatively similar results.

The next result in Figure 5.4 displays computed subglottal air pressure before m_1. It should be compared with Figure 2.2. The double peaks are qual-

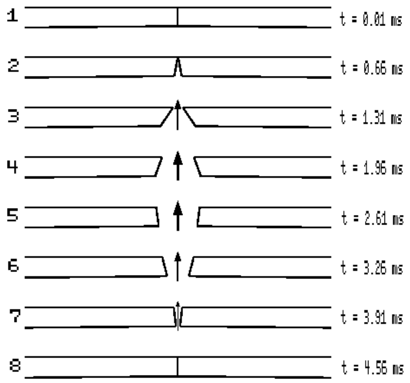

Figure 5.1: A simulated vocal fold cycle. This is similar to the figure on p. 113 of [30].

itatively similar, but there is enough difference to deserve comment. Figure 2.2 is from experimental intraglottal pressure and includes both air pressure for the open glottis and contact pressure during closure, while ours consists of only air pressure. Thus, the second peak in Figure 2.2 is gradual air pressure change in relation to contact pressure, which is their first peak.

We also simulated transglottal pressure, which is the absolute difference between the instantaneous supraglottal and sublottal pressure. To measure supraglottal pressure, we extended the domain 0.5 cm past x_2. The results are plotted in Figure 5.5. There are three peaks per cycle in this picture. The first two correspond to the vocal fold opening at x_1 and x_2, respectively. The third peak corresponds to vocal fold closure. It is similar to the p_{trans} plot in Figure 2.4.

Model robustness is demonstrated in Figure 5.6 with variable input

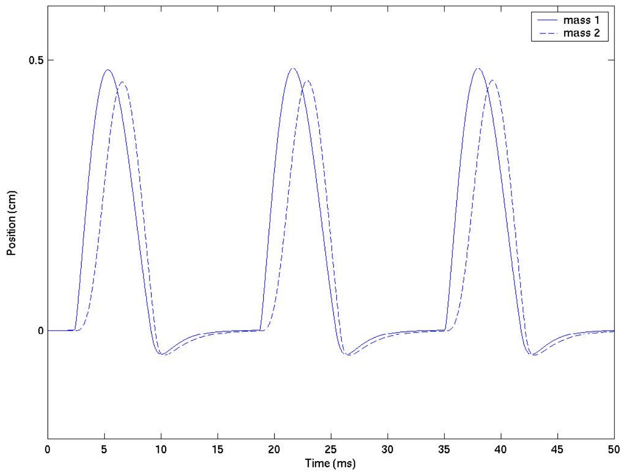

Figure 5.2: Locations of VF masses during 3 cycles of VF vibration.

pressure. Air volume velocity is displayed at the exit of the vocal folds. The three subglottal input pressures were 700 Pa, 1400 Pa, and 2100 Pa. Note that as the input pressure increases, the peak amplitudes get higher while the sides get steeper. This is in qualitative agreement with Figure 2.14(a), p. 78, in Stevens [32].

Figure 5.7 displays results of a convergence test on the numerical method under grid refinement. It consists of converging responses of air volume velocity at the exit to the vocal folds with constant input pressure.

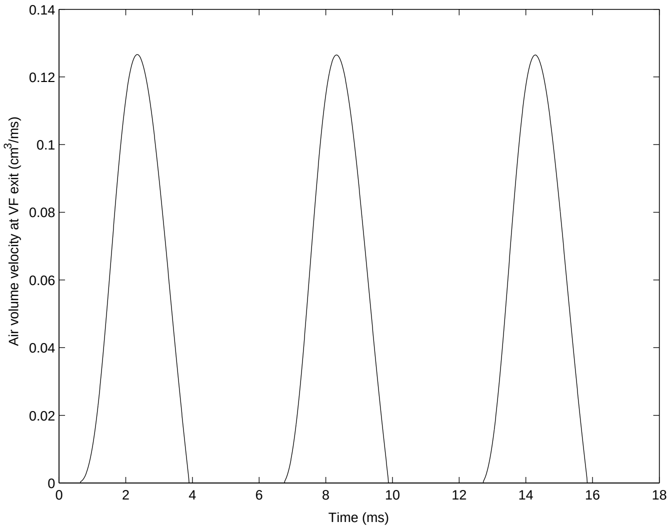

Figure 5.3: Simulated VF air volume velocity at exit of VF.

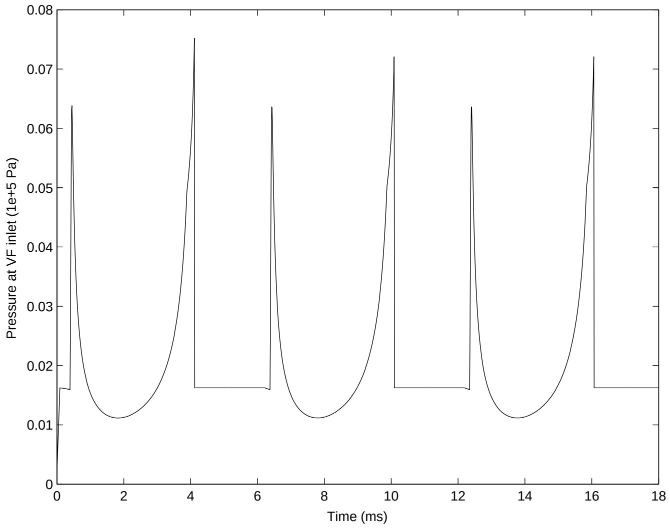

Figure 5.4: Subglottal air pressure before x_1.

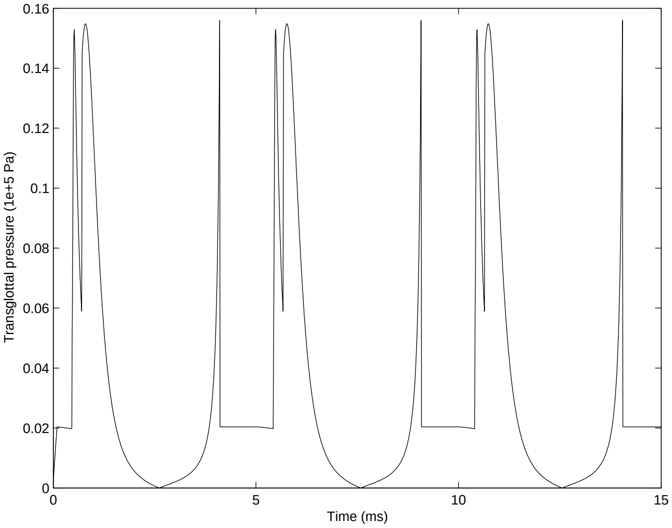

Figure 5.5: Computed transglottal pressure on an extended domain of 0.5 cm beyond x_2.

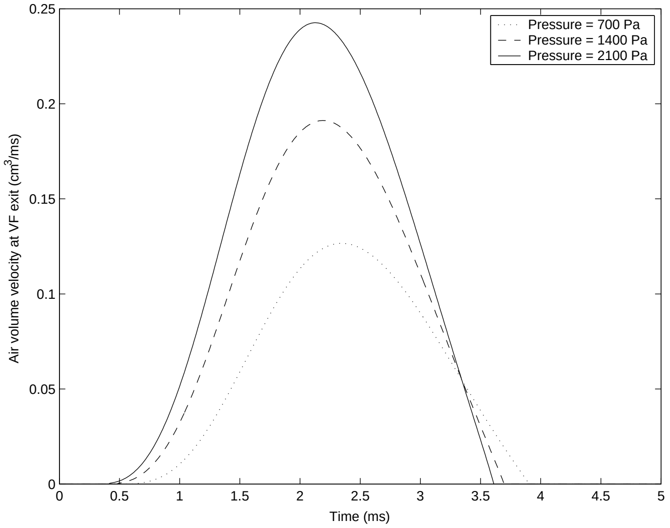

Figure 5.6: Simulated air volume velocity at x_2 for three different input subglottal pressures.

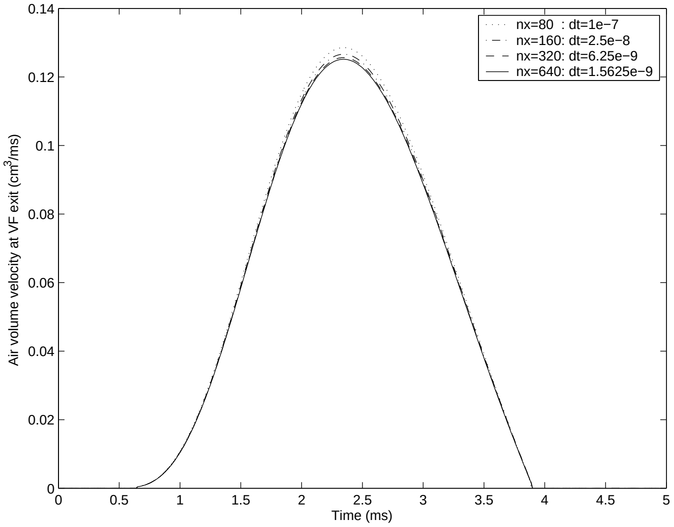

Figure 5.7: Convergence test on the numerical method under grid refinement. The plots consist of air volume velocity at the vocal fold exit.

Chapter 6

Conclusion

Part I was devoted to the description and modeling of the vocal folds. We used Bogaert's model [5] of the vocal folds, which consists of an improvement on Ishizaka and Flanagan's two-mass model [17]. The separation point was computed using an emperical formula [5, 26], and vocal fold closure was handled using an elastic collision criterion [5, 17]. The airflow equations were found using asymptotic analysis on the two-dimensional compressible Navier-Stokes equations. We derived quasi-one dimensional equations that are more accurate than Bernoulli's law, yet simpler than full two-dimensional Navier-Stokes treatments. The numerical method was described and shown to be robust and convergent. Simulation results were shown to be in qualitative agreement with many known vocal fold characteristics.

As mentioned in the chapter on numerics, future work includes numerically treating the conservation law using Gudonov's method, and then moving on to higher-order methods. It would also be beneficial to increase the complexity of the tissue from 2 mass to multi-mass, or even to a continuum. More accurate treatment of the separation point and vocal fold closure would be a fruitful avenue to explore as well.

Part II

Cochlea

Chapter 7

Introduction

Table 7.1: Part II: List of abbreviations

BM	Basilar membrane
TM	Tectorial membrane
CP	Characteristic place
CF	Characteristic frequency
SS	Steady state
TD	Time domain
OC	Organ of Corti
PA	Pressure amplitude
IHC	Inner hair cell
OHC	Outer hair cell
FFT	Fast Fourier Transform
SPL	Sound pressure level

Mathematical modeling of hearing has benefited us in many ways. Hearing aids and cochlear implants are examples in the medical field. We can diagnose many ailments of hearing through our understanding of the hearing process. This understanding has come from experiments, but much of it has come from theory as well due to the difficulty in extracting data from the hearing process. The ear is highly susceptibility to damage and is a very sensitive organ. Audio compression, as well as noise suppression, are other benefits that we experience through our existing modeling efforts. However, there is room for improvement, especially in the realm of nonlinear cochlear models. Nonlinear models are able to capture many hearing effects, such as suppression and difference tones, that linear models alone cannot resolve. Improvements to noise suppression and audio compression are possible. The end result of Part II will be a nonlinear nonlocal cochlear model based on Neely and Kim's [25] active linear model.

We start by asking the question "What is hearing?" Most of us have five senses to gather information and experience from the world. Hearing is the sense capable of measuring pressure differences in the fluid air medium that we live. How does the ear sense these minute differences? Exactly how sensitive is the ear? What is the ear capable of hearing? There are two natural approaches to hearing, each answering these questions in their own ways. The first approach is based on working to understand the *mechanics* of hearing by concentrating on the organ of the ear. The second approach is on working to understand the *perception* of hearing and takes place in the higher auditory centers of the brain. This paper is concerned with the first approach.

Let us look now at how sensitive the ear is and what the ear is capable of hearing. There are three measurements we will look at, with an illustrative analogy given for each from [6].

- Frequency selectivity

  The first is frequency selectivity. Consider two adjacent keys on a piano. They have a relative frequency difference of 6%. What frequency resolution does the ear have? The ear is capable of distinguishing two frequencies that have a relative frequency difference of 0.2% (see Figure 1.1 in [6]).

- Sensitivity

How sensitive is the ear to pressure fluctuations? In chapter 8, we will describe hair cells in the ear that transduce mechanical motion to electrical impulses. Take one of these hair cells and scale it to the height of the Sears tower. The hair cell is capable of detecting motion on subatomic scales, equivalent to the Sears tower moving a mere 5 cm (see Figure 1.2 in [6]).

- Dynamic range

  Dynamic range refers to the ear's ability to detect sounds over an intensity range spanning over a million fold change in energy. This is equivalent to the ratio of the weights of five elephants to one mouse (see Figure 1.3 in [6]).

These examples are meant to stress the impressive abilities of the ear to detect sound and highlight the challenges to modeling such a system.

Beginning measurements were done by Békésy [36] on dead cochlea from cadavers. His measurements were unable to explain the above three properties of the cochlea. Much later, it was discovered that the live cochlea has an active feedback process that enables it to achieve the sensitivity and dynamic range mentioned above. For very loud sounds, the ear acts like a dead cochlea, which we refer to as the *passive* cochlea. For soft to midrange intensities, the ear is supplying amplification to the sound and thus we refer to this as the *active* ("live") cochlea.

We begin in Chapter 8 by discussing the biology of the outer, middle and inner ear. We discuss the difference between a "dead" cochlea (passive) and a "live" cochlea (active). An active cochlea receives feedback from the higher auditory processing centers to help amplify and tune the cochlea. We then display current measurements from the literature that we will compare with our model results.

Chapter 9 begins our modeling efforts by looking first at the linear passive cochlea. The cochlea is filled with an incompressible fluid. Using fluid mechanics, we derive the well-known box model from the full Navier-Stokes equations and then solve this model using separation of variables, creating a fluid functional that adds an inertial term to the cochlear partition (i.e. tissue). The cochlear partition consists mainly of what is called the basilar membrane (BM), which is tuned from high to low frequencies in a continuous decreasing manner from the entrance of the cochlea to the apex. The basilar membrane acts as a Fourier transform by picking out each frequency component's vibrational waveform from the incoming sound. We start with a simple model of the cochlear partition consisting of just the basilar membrane to demonstrate the coupling between the cochlear partition and the cochlear fluid. We then move to a more complicated cochlear partition model based on Neely and Kim [25] consisting of the basilar membrane with a second resonating membrane called the tectorial membrane (TM), which we call the TM/BM model. For completeness, we end the chapter with a discussion of the one-dimensional long-wave approximation.

Chapter 10 discusses the middle ear model and some of the subtleties involved in its coupling with the cochlea. In particular, we work in the time domain to prepare for the nonlinear treatment of the cochlea. In so doing, we discover a dispersive instability near the apex of the cochlea that is not present in the frequency domain model. The dispersive instability is present when a frequency domain middle ear is coupled with the time domain cochlear model, but disappears when a time domain middle ear is coupled with the time domain cochlear model. This chapter introduces concepts from engineering vibration analysis in order to better understand how the middle ear acts as a vibration amplifier and transient absorber for the cochlea.

Chapter 11 moves to the linear active case using Neely and Kim's well-known model [25]. We generalize their active gain to a nonlinear nonlocal functional of BM displacement, which completes our model setup. We move to the numerical treatment of this system and discuss some simplifications and efficiencies in its implementation on the computer.

Chapter 12 displays our model results, reproducing a number of nonlinear effects present in the cochlea, such as combination difference tones and suppression. We compare our results with the data in Chapter 8. We complete the results section with the introduction of a cochlear response spectrogram (cochleogram) that is shown to resolve more transient micro-structures than the FFT spectrogram.

Chapter 13 concludes Part II on the cochlea and discusses future work. Much of the work in Part II is given in [20].

Chapter 8

Cochlear Physiology

This chapter covers first the anatomy of the ear and then moves to the measurements and experiments that have been performed on the ear. The section on measurements will give us results to compare the feasibility of our modeling efforts and will shed some light on to how the biology performs its function. The interplay between modeling and experiment will be stressed.

Figure 8.1 shows the three sections of the ear: the outer ear, middle ear and inner ear. The majority of this paper will look at the inner ear, where the complexity of hearing in these sections takes place. However, the interplay between the middle and inner ear is important both biologically and numerically in our models, and therefore will be discussed at length throughout.

8.1 Outer Ear

Let's follow a sound wave through the ear to introduce the different types of processes that occur. A pressure wave begins in air and travels at the speed of sound to the outer ear. Here, the pinna localizes and channels the sound to the ear canal, where it travels unaffected to the tympanic membrane, or eardrum.

The pinna is used to focus the sound waves and reduce the amount of acoustic reflection with the skull. It is a way to externally locate sound sources.

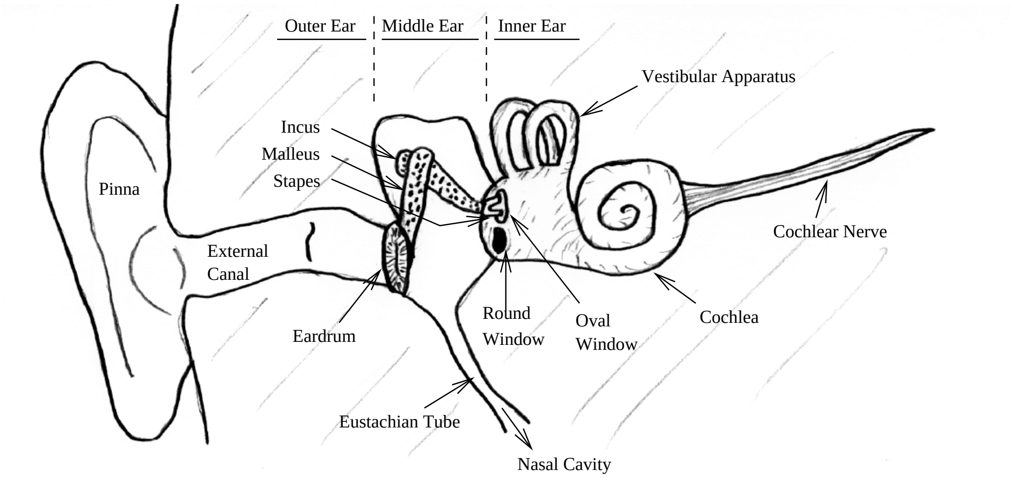

Figure 8.1: Schematic of the outer, middle and inner ear (Reconstructed from [28]).

This can be seen in those lucky animals that have the ability to change the shape of their pinna towards the sound of interest in order to better localize the sound.

Note that the middle and inner ear are recessed within the skull. The ear canal provides the separation needed to protect the delicate and sensitive eardrum and cochlea. The ear canal simply transfers the air waves from the pinna to the cochlea with very little transformation of energy.

8.2 Middle Ear

Continuing with the sound wave, at the end of the ear canal the eardrum is set in motion. The eardrum is connected to a series of small bones in the middle ear (ossicular bones) that act as levers to transduce the sound from air to fluid in the cochlea. The ossicular bone connected to the inner ear (cochlea) via the oval window is known as the stapes. It vibrates, creating a standing pressure wave in the cochlear fluid.

The area of the eardrum is larger than the area of the oval window, giving an amplification to the signal. This is demonstrated in Chapter 10 on the middle ear.

8.3 Inner Ear

Again returning to our sound wave, the vibration of the stapes creates a standing pressure wave in the cochlear fluid. Immersed in the cochlear fluid is a membrane, called the basilar membrane (BM), that is stretched along the entire length of the cochlea. The BM is set in motion by the fluid pressure difference across the membrane. Each point on the BM is tuned to resonate at varying frequencies: high frequencies (about 20 kHz) at the base near the stapes to low frequencies (about 20 Hz) near the apex. The motion of the BM is in the form of a traveling wave that travels along the BM from base to apex. At the apex of the cochlea is the helicotrema, which acts as a pressure release for very low sounds (out of the range of hearing).

Perched on top of the BM is a relatively tiny system known as the Organ of Corti (OC). This is depicted in Figure 8.2. As the BM vibrates, the inner hair cells (IHC) anchored to the BM are shear displaced, creating a neural pulse that is sent to the brain. This is the information that is decoded into the perception of hearing.

8.3.1 Active Feedback

So far, the above description describes what is known as a passive cochlea, where we have assumed that all energy absorbed by the BM comes from the input at the stapes. However, this does not account for the extreme sensitivity and dynamic range of the cochlea. There is consensus in the cochlear community that there is an active process occurring within the system which effectively *adds* energy to the BM motion. This process has been shown to take place in the outer hair cells (OHC) (see Figure 8.3). There are nerve fibers conveying information to the OHCs from the higher neural auditory centers, showing that the cochlea is receiving *feedback*.

The exact method of active feedback is under constant debate from the cochlear community. The viewpoint that we take in this paper is mostly towards signal processing, so the modeling effort in the hair cells does not contain much depth. We assume that some active process is occurring which changes the BM tuning properties, and that is all that we need for this treatise.

We will look at what we call the *macromechanics* of the cochlea, consisting of the cochlear shape, fluid chambers and vibrating basilar membrane. As was mentioned in the previous discussion, perched on top of the basilar membrane is the relatively tiny Organ of Corti, which contains the tectorial membrane (TM) and hair cells. This is often referred to as the *micromechanics* of the cochlea.

8.3.2 Macromechanics

A cross-section of the cochlea is shown in Figure 8.2. There are three fluid chambers in the cochlea: the scala vestibuli, scala media and scala tympani. The Reissner's membrane separates the scala vestibuli from the scala media, while the basilar membrane separates the scala media from the scala tympani. The upper and lower chambers contain perilymph, while the middle chamber housing the Organ of Corti contains endolymph. Endolymph has a

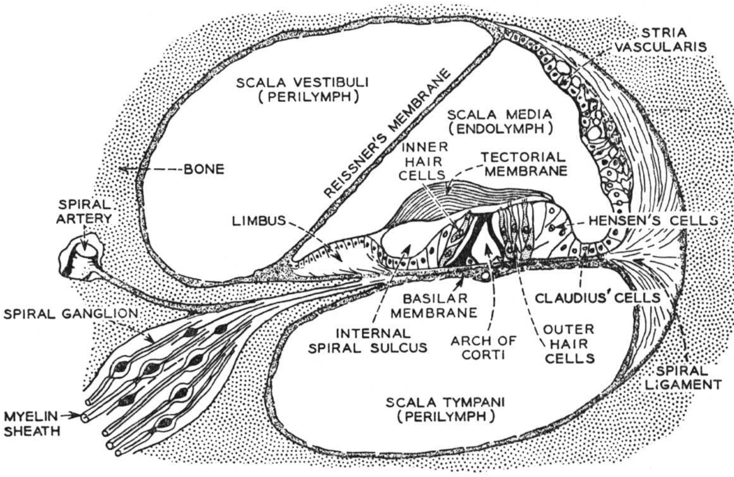

Figure 8.2: Macromechanical cross-section of the cochlea, showing in particular the three fluid chambers and basilar membrane, with the Organ of Corti (OC) perched on top [2]. (From Physiology of the Ear 2nd edition by JAHN. © 2001. Reprinted with permission of Delmar Learning, a division of Thomson Learning: www.thomsonrights.com. Fax 800 730-2215.)

high potassium content and is necessary for the mechanical-to-electrical transduction that takes place in the Organ of Corti. The potassium is created in the stria vascularis, which is innervated with blood vessels. It is kept at a distance from the sensitive Organ of Corti due to the relatively loud noise created by blood flow. The Reissner's membrane has little affect on the fluid flow in the cochlea and is thus commonly ignored (it is only two cell layers thick). In our modeling efforts, therefore, we will model the cochlea as having two chambers: the scala vestibuli and the scala tympani.

The basilar membrane is tuned to resonate at different frequencies, from higher frequencies near the stapes to lower frequencies near the apex. Fluid pressure difference between the scala vesitbuli and scala tympani drives the BM motion. Note that while the dimensions of the cochlea are in centimeters,

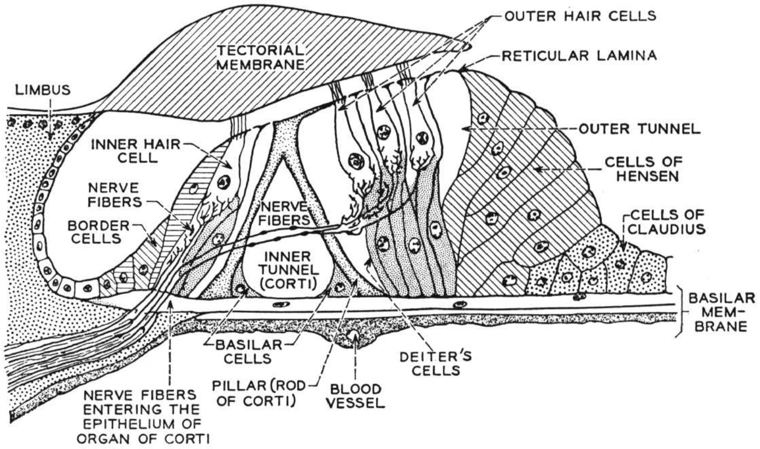

Figure 8.3: A magnified view of the Organ of Corti [2]. (From Physiology of the Ear 2nd edition by JAHN. © 2001. Reprinted with permission of Delmar Learning, a division of Thomson Learning: www.thomsonrights.com. Fax 800 730-2215.)

the vertical BM motion is in nanometers, or atomic scales.

8.3.3 Micromechanics

Figure 8.3 shows a magnified view of the OC. The OC houses the hair cells, both the inner and outer hair cells, and is where mechanical vibration is transduced to electrical impulses that are sent to the higher auditory processing centers. The inner hair cells house afferent nerves (i.e. away from the ear), while the outer hair cells have mostly efferent nerves (to the ear), as well as afferent nerves. This points to two-way communication with the outer hair cells and the brain, showing a feedback process occurring.

The outer hair cells are imbedded in the tectorial membrane. Motion of the basilar membrane causes motion of the tectorial membrane, giving a shear displacement of the hair cells. An electrical impulse is then released when the hair cells are displaced in a particular direction. The tectorial membrane is thought to be a second-filter for mechanical vibration in the ear due to its coupling with the basilar membrane through the outer hair cells. We implement this idea in our modeling efforts.

8.4 Measurements

Any attempt at quantifying a human experience, in particular through a sensory organ such as the ear, runs into the fact that there is a difference between reality and our perception of reality. We can construct operational definitions of things in our surroundings, such as temperature, location, speed, light, etc. However, they are only actions that we use to create a description of an experience. This is no different in the case of hearing. The prime example is frequency vs. pitch. We can measure frequency in the world, but pitch is our *experience* of sound. Thus, a division of study in hearing is in the mechanical measurements and the psychoacoustical measurements. We will concentrate in this treatise on the mechanical measurements. So the ultimate question is what do we measure in the cochlea to quantify hearing? Based on our past discussion, we can measure the vibration of the BM, either its displacement or velocity. We can also measure the neural firing rates of the auditory neurons. What this leaves out is *how* we are to measure these. The cochlea is imbedded deep in the skull next to the brain and completely surrounded by bone. Even when we get to the cochlea, the quantities we are measuring are on the order of atomic scales, and the sensitivity of the cochlea makes it very susceptible to damage. Despite the challenges pointed at above, however, measurements have been made along the cochlea [29, 36]. This section will discuss the typical measuring techniques and will serve as the templates to compare our model results to.

8.4.1 Units

It is quite fascinating when one realizes that the cochlea can detect mechanical movements on the atomic scale. So while the cochlear length is on the order of centimeters (10^{-2} meters), the BM response is on the order of only nanometers (10^{-9} meters). It is important, therefore, when viewing figures to pay particular attention to the quantity being displayed and its correct units. This applies similarly to frequency and time. As mentioned previously, the frequency range of the ear can span 20 Hz to 20 kHz, or three orders of magnitude (10 octaves). In our figures, we will usually display frequency in the units of kHz in a log scale, due to the dynamic range of frequency detection. Nearly every measurable quantity in the ear spans a large dynamic range. This leads us to the use of the decibel (dB) in representing quantities. The dB can be used for any unit. It should first be noted that dB is defined *relative* to a landmark unit. For example, for pressure in acoustics we define the dB SPL (sound pressure level) as

$$\text{dB SPL} = 20\log_{10}\left(\frac{p}{p_0}\right)$$

where $p_0 = 20\ \mu\text{Pa}$. p_0 represents an experimentally verified input pressure amplitude (PA) of a single 1-3 kHz tone at the threshold of hearing for a normal healthy ear. Thus, a range of $0 - 120$ dB SPL spans six orders of magnitude in pressure relative to the pressure at the threshold of hearing. For an arbitrary unit x, we define dB relative to R by

$$\text{dB re R} = 20\log_{10}\left(\frac{x}{R}\right)$$

When x is displacement, then typcially we still refer to it as dB SPL, where now we have the landmark unit $R = 1$ nm. This is the BM response displacement for an input 1-3 kHz tone at the threshold of hearing. Similarly, when x is velocity, the landmark unit is $R = 100$ μm/s.

8.4.2 Global View

The easiest measurement to grasp is the global panoramic view of the BM vibration. This consists very simply of giving the cochlea a simple harmonic input and observing the BM vibration all along the length of the cochlea. This is known as the *BM response* and is illustrated in the left plot of Figure 8.4 for a single 2 kHz tone at 30 dB SPL (in this case the magnitude of the BM response). Let us denote the magnitude of the steady state BM response by $U(x, \omega_0, A_0)$, where x is BM location and the input is a single harmonic input with SPL A_0.

8.4.3 Local Views

While the BM response is the simplest description of cochlear vibration, it is by no means the easiest in terms of measurement. The typical means of measuring cochlear vibration is to place a probe at a specific location along the basilar membrane and record the magnitude of the response at that position for varying frequencies and intensities. This is a *local* measurement. The types of curves produced by taking local measurements are called iso-contour curves, or frequency tuning curves. Let's look at two specific examples.

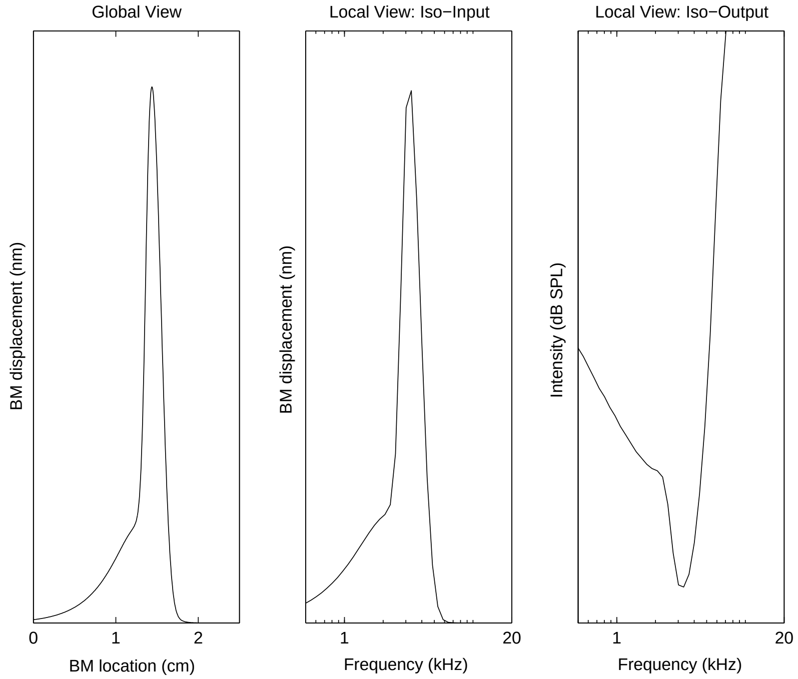

Figure 8.4: The left plot is a global view of the BM. It is the absolute BM steady state response for a 2 kHz tone at 30 dB SPL. The middle plot is a local view iso-input curve at BM location $x = 1.2$ cm. The right plot is a local view iso-output curve at the same BM location.

8.4.3.1 Iso-Input

Iso-input curves are produced by measuring the magnitude of the steady state response (BM displacement/velocity) at a specific location on the BM while varying either the frequency and/or the SPL. An example of an iso-input curve where the SPL is fixed and we vary the frequency is given in the middle plot of Figure 8.4. We fix a BM location x_0 and feed the cochlea a harmonic input with frequency ω and SPL A_0. Thus, an iso-input curve is given by $U(x_0, \omega, A_0)$. Mathematically, it is the same function as the global view except we interchange the variable between x and ω.

8.4.3.2 Iso-Output

Iso-output curves (or iso-response curves) are found by plotting the sound level required to achieve a certain maximum BM response (displacement or velocity) at a specific location for varying frequencies. When we are measuring displacement/velocity, the curves are called iso-displacement/iso-velocity curves, respectively. An example of an iso-displacement curve is given in the right plot of Figure 8.4. Formally, we can denote an iso-displacement curve at a point x_0 on the BM as a sound pressure level function $A(\omega)$ such that

$$U(x_0, \omega, A(\omega)) = R \tag{8.1}$$

where R is the maximum BM response in nanometers. Typical values for R include 1 nm for BM displacement and 100μm/s for BM velocity. These values are based on the threshold of hearing for young healthy cochlea.

8.4.4 Characteristic Maps

Each location on the basilar membrane is tuned mechanically to resonate at a specific frequency. However, the resonant place for a particular frequency is not at the location of maximal response. The location of maximal response occurs basal to the resonant point due to damping. On the BM, the resonant place is where the BM displacement first drops to zero, denoting the start of the cutoff region. The location we are interested in is not the resonant location, but the place of maximal response. The relation of input frequency to location of maximal response is called the characteristic frequency-to-place map, denoted $\mathcal{X}(\omega)$. The inverse of this map, called the characteristic place-to-frequency map and denoted by $\mathcal{W}(x) \equiv \mathcal{X}^{-1}(x)$, relates BM location to frequency that gives the maximal response. It is important to realize that how we define either of these maps needs to be consistent with the above description, i.e. $\mathcal{W}$ and $\mathcal{X}$ need to be inverses of each other. The following questions highlight the issues:

- Are we defining the map based on experiment or based on an existing active model? How do we define the map so that it is independent of which we choose?
- In a nonlinear active cochlea, the maximal response depends on the input SPL (the well known half-octave shift [29]). Which SPL do we use? Can we define a map that is "independent" of SPL?

If we define the map based on experiment, we have a "local" view of the cochlea and want to use iso-contour curves. If we define the map based on an active model, we have a "global" view of the cochlea and can use BM displacement responses. We could just decide to define $\mathcal{W}$ or $\mathcal{X}$ using one of these in both experiment and modeling, but it is very important that there be consistency between them for comparison purposes.

Let us be a little more precise in our discussion. We feed the cochlea a harmonic input with frequency ω_0 and SPL A_0. The magnitude of the BM response is given by $U(x, \omega_0, A_0)$, and thus we define $\mathcal{X}(\omega_0)$ by

$$U(\mathcal{X}(\omega_0), \omega_0, A_0) \equiv \max_{x \in [0,L]} U(x, \omega_0, A_0) \tag{8.2}$$

This is the "global" view and is useful when we have an actual model.

There are two options for the definition of $\mathcal{W}(x)$. Fixing a BM location x_0 and SPL A_0, we define $\mathcal{W}(x_0)$ by

$$U(x_0, \mathcal{W}(x_0), A_0) \equiv \max_{\omega} U(x_0, \omega, A_0) \tag{8.3}$$

The second option is based on iso-displacement curves and is defined by

$$U(x_0, \omega, A(\omega)) = 1 \Longrightarrow A(\mathcal{W}(x_0)) \equiv \min_{\omega} A(\omega) \tag{8.4}$$

Both of these options are "local" definitions and are useful in experimental work.

Note that definitions (8.2) and (8.3) are found using the same procedure, namely finding steady state responses over a range of frequencies and maximizing either over the entire BM response or maximizing over all frequencies at a particular BM location. Definition (8.4) is vastly different (on the surface). This definition, however, is very useful since it is *independent of input SPL*, whereas (8.2) and (8.3) depend on the input SPL. If we use definitions (8.2) or (8.3), which SPL do we use? How do we choose it to be consistent with definition (8.4)? To answer these questions, we will first look at the linear cochlea.

8.4.4.1 Linear cochlea

In the linear cochlea, definitions (8.2) and (8.3) are *indepedent of SPL*. This is seen by (denoting A as pressure amplitude instead of SPL)

$$U(x, \omega, A) = H(x, \omega)A \Longrightarrow \frac{U}{A} = H(x, \omega) \tag{8.5}$$

where $H(x, \omega)$ is called the *frequency response* of the system and is in effect the ratio of output to input.

To show (8.3) and (8.4) are equivalent, we note that

$$\begin{aligned} U(x_0, \omega, A(\omega)) = 1 &\iff H(x_0, \omega)A(\omega) = 1 \\ &\iff H(x_0, \omega) = \frac{1}{A(\omega)} \\ &\iff \max_{\omega} H(x_0, \omega) = \frac{1}{\min_{\omega} A(\omega)} \end{aligned}$$

Thus, the maximum and minimum are attained at the same point $\omega_0 \equiv \mathcal{W}(x)$, showing both definitions based on local views are equivalent.

8.4.4.2 Nonlinear cochlea

We showed above that for a linear system, (8.4) $\equiv$ (8.3). For the nonlinear active case, each definition depends either on the SPL or the response. Thus, we must make sure that for (8.4), we choose a maximum response R where the system is nearly linear with near full active gain. For (8.2), we must choose a corresponding SPL where the system is, again, nearly linear with near full active gain.

This is illustrated as follows. We will start by defining the map $\mathcal{W}(x_0) = \omega_0$ using definition (8.4). We will choose a sufficiently small maximum response R so that the ear has full active gain (typically 1 nm). Thus, we have $U(x_0, \omega, A(\omega)) = R$ and $A(\omega_0) = \min_\omega A(\omega)$. Let P_0 denote the point $(x_0, \omega_0, A(\omega_0))$. We have

$$\begin{aligned} U(x_0, \omega, A(\omega)) = R \quad &\Longleftrightarrow \quad \frac{\partial U}{\partial \omega} + \frac{\partial U}{\partial A} \cdot \frac{dA}{d\omega} = 0 \\ &\Longleftrightarrow \quad \left.\frac{\partial U}{\partial \omega}\right|_{P_0} + \left.\frac{\partial U}{\partial A}\right|_{P_0} \cdot \left.\frac{dA}{d\omega}\right|_{\omega_0} = 0 \end{aligned}$$

Assuming $A(\omega)$ is a smooth function, since ω_0 is a minimum of A, we have

$$\left.\frac{dA}{d\omega}\right|_{\omega_0} = 0$$

which gives

$$\frac{\partial U}{\partial \omega}(x_0, \mathcal{W}(x_0), A(\omega_0)) = 0$$

We were thus able to show that $(x_0, \mathcal{W}(x_0), A(\omega_0))$ is a critical point of the iso-input curve. If we assume this critical point is a maximum, then we have the equivalence of (8.4) and (8.3) *as long as the SPL used in (8.3) is* $A(\omega_0)$. The problem is that different iso-output curves will give different SPLs, but our iso-input curve is defined using only one SPL. To remedy this, if the maximum BM response R used in the definition of the iso-output curve above is chosen small enough so that the resulting range of SPLs $[A_0, A_1]$ is such that the ear is *nearly* linear with *near* full active gain, and we choose an SPL $A \leq A_0$ for our definition of the iso-input curve, then the characteristic place-to-frequency maps will be *nearly* equivalent.

To summarize, for a given BM location x, there is a *characteristic frequency* (CF) where x is the location of maximal response to a harmonic input with frequency given by the CF and SPL such that the ear is fully active. Conversely, given a frequency ω in the range of hearing, there is a *characteristic place* (CP) on the BM that satisfies the same criteria.

8.4.5 Nonlinearities

8.4.5.1 Nonlinear Compression (Dynamic Range)

The dynamic range of the ear is demonstrated experimentally through the use of an iso-intensity curve, given in Figure 8.5. This is also known as a sensitivity curve, for it displays the ratio of output to input. The iso-intensity curve on the left is an iso-input curve, i.e. a local view, where a probe was placed at the BM location with characteristic frequency 10 kHz. Similar to equation (8.5), we define

$$\frac{U}{A} = H(x, \omega, A) \tag{8.6}$$

where again H is the frequency response of the system. This time, however, H depends on the input SPL A, since the system is nonlinear. Because of this, the iso-intensity curves will be different for different input SPL, and in

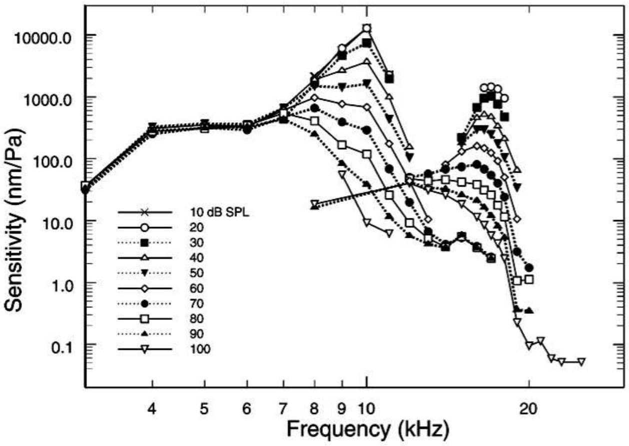

Figure 8.5: This is a collection of sensitivity plots at a BM location with characteristic frequency 10 kHz. The input SPL was varied from soft to loud to demonstrate the detuning of the cochlea (reproduced Fig. 5 on p. 1311 from [29], with permission).

fact will be *more sensitive* for lower SPL, as this is when the cochlea gives the most active gain. Looking at the picture, we see that for low SPL, the ear is finely tuned and its peak is at the characteristic frequency. However, as the SPL increases, the sensitivity decreases and the peak moves to lower frequencies. This is called the *half-octave shift*, since the difference between the peak frequencies for the fully passive vs. fully active cochlea is around 1/2-octave.

The plot in Figure 8.6 shows sensitivity plots at a particular BM location with an impulse input, or click. The BM response consists of velocity instead of displacement. Again we see that as the amplitude of the input is increased, the cochlea becomes less sensitive. This nonlinear behaviour allows the ear to cover a wider range of inputs.

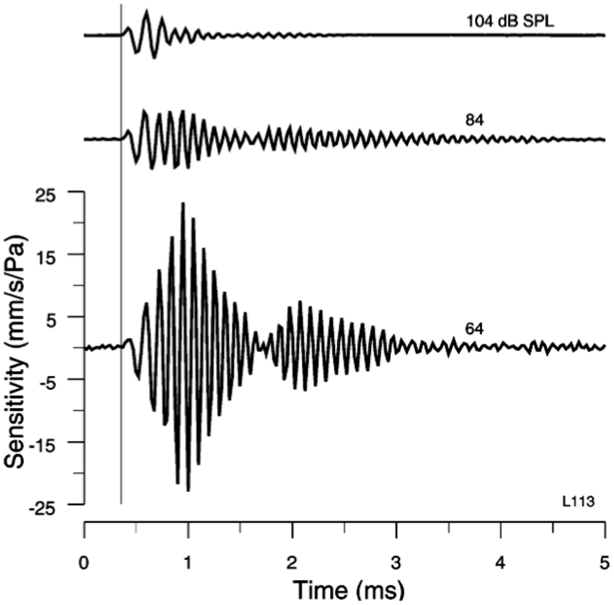

Figure 8.6: A collection of sensitivity plots for an impulse input, or click (reproduced Fig. 9 on p. 1314 from [29], with permission).

8.4.5.2 Combination Difference Tones

A nonlinear cochlea will create what are known as combination difference tones when more than one tone are presented to the ear. For example, if two frequencies f_1 and f_2 are put into the ear, $nf_1 \pm mf_2$ will be created at varying intensities, where n and m are nonnegative integers. The cubic difference tone, denoted $f = 2f_1 - f_2$, where $f_1 < f_2$, is the most prominent. In Figure 8.7, two tones f_1 and f_2 are input into the ear such that the cubic difference tone is given by $2f_1 - f_2 = 7.5$ kHz. A probe is placed on the BM at the characteristic place for 7.5 kHz and the FFT is performed on the BM response waveform. We see a collection of difference tones, with the cubic difference tone indeed being the most prominent tone.

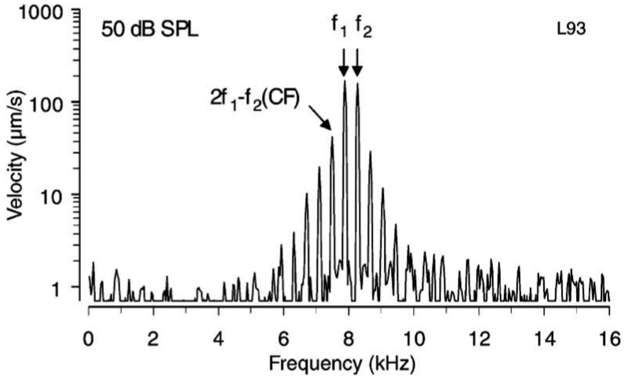

Figure 8.7: Cubic difference tones are displayed for two tones at 50 dB SPL so that $2f_1 - f_2 = \text{CF}$ (7.5 kHz). (reproduced Fig. 17 on p. 1330 from [29], with permission)

8.4.5.3 Suppression & Masking

Masking occurs when the presence of one frequency in the BM response masks, or suppresses, the BM response of another frequency. Audio compression, such as MP3, uses this characteristic of the ear to compress signals [27]. Figure 8.8 contains iso-velocity contours at a location on the BM that has a characteristic frequency of 10 kHz. The dotted line is the tuning curve without added masking signals. Note that the minimum of the curve is where it should be at 10 kHz. The solid curves represent the same experiment, but now with the presence of masking tones, consisting of a 500 Hz tone in one curve and a 12 kHz tone in another. Note that the location has been "detuned" in that it is less sensitive and the minimum occurs at a lower frequency. Also note that the lower suppressor tone (500 Hz) has more of a masking effect than the closer 12 kHz tone. This is called the upward spread of masking and shows that low-side suppression is stronger than high-side suppression.

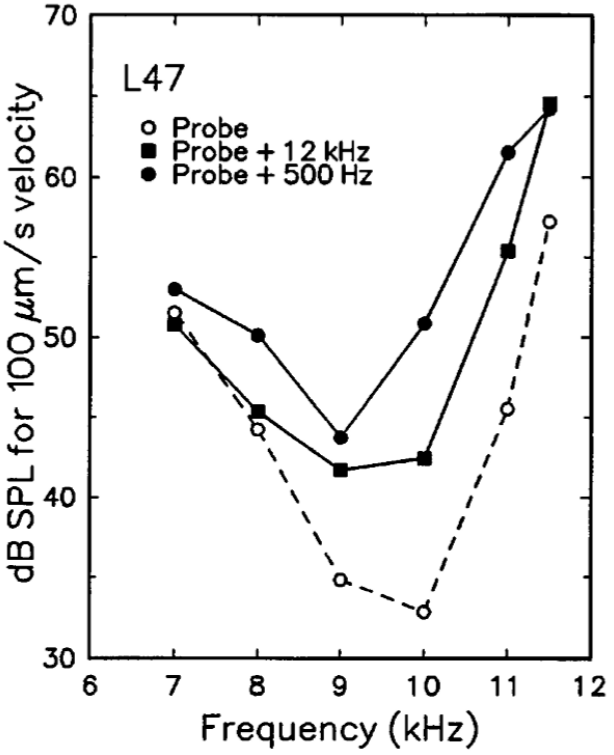

Figure 8.8: Iso-velocity plots (iso-output at 100 μm/s) for BM responses at the characteristic place for 10 kHz. The two solid lines are tones in the presence of high-side (12 kHz) and low-side (500 Hz) suppressors, both input at 70 dB SPL. (reproduced Fig. 16 on p. 1327 from [29], with permission)

Chapter 9

The Passive Cochlea

We begin our modeling efforts by looking first at what is known as the macromechanical structure of the cochlea, which consists of the overall cochlear shape, fluid structure and cochlear partition, discussing each in the order just mentioned. For simplicity, we will concentrate our efforts on the passive, or dead cochlea.

9.1 The Box Model (Three-Dimensional Model)

As was seen in the previous chapter, the cochlea is spiral shaped. Does its shape influence the mechanics of internal motion? The cochlear dimensions are such that the shape does not effect the mechanics to a large extent, and thus we will ignore this spiral structure by using what is known as the box-model of the cochlea. See Figure 9.1 for an illustration. We will at first assume a three dimensional cochlea, reducing the dimension as we go to simplify the

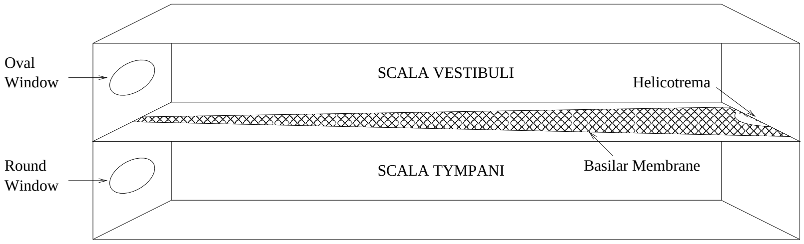

Figure 9.1: Schematic of the 3D box cochlear model.

modeling.

Remark 9.1.1. The dimension of the model refers to the fluid dimension. The cochlear partition is in all cases of co-dimension one to the fluid. The special case is the one-dimensional long-wave model of the cochlea. In this instance, the cochlear partition is also one-dimensional. The one-dimensional model is found by "factoring out" the second dimension and then moving to an asymptotic limit in the wavenumber of the dispersion relation. This is discussed in the last section of this chapter.

9.1.1 The Navier-Stokes Equations: Simplifications

The following analysis follows that performed by Baker in [4]. We start with the Navier-Stokes equations for incompressible flow in three-dimensions:

$$-\nabla p + \mu \triangle \vec{v} + \rho g \vec{k} = \rho(\vec{v}_t + \vec{v} \cdot \nabla \vec{v})$$

The cochlear fluid is nearly inviscid, and thus we have $\mu = 0$, giving

$$-\nabla p + \rho g \vec{k} = \rho(\vec{v}_t + \vec{v} \cdot \nabla \vec{v})$$

9.1.1.1 Linearization

The dimension of cochlear vibrations (nanometers) relative to cochlear dimensions (millimeters) allows us to drop the convective term above and effectively linearize the equation. The inertia term is given by

$$\rho(\vec{v}_t + \vec{v} \cdot \nabla \vec{v}) \tag{9.1}$$

We will consider a standing wave, letting $\vec{v}(x,t) = \vec{v}(x)e^{i\omega t}$. We are interested in fluid displacement, so we substitute $\vec{v} = i\omega \vec{u}$ in (9.1) to obtain

$$-\rho\omega^2(\vec{u} + \vec{u} \cdot \nabla \vec{u})$$

We now nondimensionalize by making the following change of variables

$$\begin{aligned} \lambda\hat{\vec{u}} &\to \vec{u} \\ \frac{1}{L}\nabla_\xi &\to \nabla \end{aligned}$$

where $\xi = x/L$, ξ and $\vec{u}$ unit length. This gives

$$-\rho\lambda\omega^2(\hat{\vec{u}} + \frac{\lambda}{L}\hat{\vec{u}} \cdot \nabla_\xi \hat{\vec{u}})$$

Since λ is on the order of nanometers and L is on the order of millimeters, we can effectively ignore the convective term.

We will ignore the effects of gravity as well, thus considering the problem

$$-\nabla p = -\rho\omega^2\vec{u} \tag{9.2}$$

9.1.1.2 Laplace's equation

We have the incompressibility condition

$$\nabla \cdot \vec{v} = 0 \Rightarrow \nabla \cdot \vec{u} = 0$$

Taking divergence of both sides of equation (9.2), we finally arrive at Laplaces equation for the pressure

$$\triangle p = 0 \tag{9.3}$$

9.1.2 Boundary Conditions (Two-Dimensional Model)

The analysis in the previous sections applies to the two-dimensional fluid case as well, and thus for simplicity we will concentrate from this point on the two-dimensional fluid model (with a one-dimensional cochlear partition

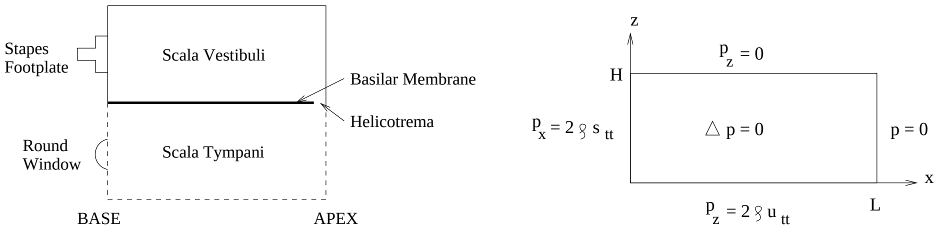

Figure 9.2: The figure on the left is a schematic of the cochlea, while the figure on the right represents the upper chamber with the macromechanical equations and boundary conditions.

model). The two-dimensional model setup and structure is illustrated in Figure 9.2.

The stapes boundary condition is given by

$$p_x(0, z, t) = 2\rho s_{tt}(t) \tag{9.4}$$

where ρ is fluid density, $p(x, z, t)$ is fluid pressure and $s(t)$ is stapes displacement. Equation (9.4) is $F = ma$, where the force F is pressure flux, mass m is fluid density, and acceleration a is stapes acceleration. It is important to note the dependence on stapes acceleration, NOT stapes displacement. This in effect turns the middle ear from a low-pass filter to a high-pass filter. More discussion on middle ear effects is given in Chapter 10.

For the apical boundary condition, we will consider for the moment a Dirichlet boundary condition

$$p(L, z, t) = 0 \tag{9.5}$$

There is some revealing analysis when one considers a Neumann condition in terms of controlling the dispersive instability at the apex (see [38] and the discussion in Chapter 10). In the steady state case (see section 9.2.1), the

Dirichlet and Neumann conditions do not differ. Neely & Kim [25] consider an absorbing boundary condition in their paper, but again the difference is negligible in the steady state case.

The upper wall is rigid, and thus the boundary condition is given by

$$p_z(x, H, t) = 0 \tag{9.6}$$

The basilar membrane condition is given by

$$p_z(x, 0, t) = 2\rho u_{tt} \tag{9.7}$$

where $u(x, t)$ denotes basilar membrane displacement.

9.1.3 Separation of Variables

Combining (9.3) with the boundary conditions (9.4)–(9.7), we get

$$\begin{cases} \triangle p(x, z, t) = 0 \\ p_x(0, z, t) = 2\rho s_{tt}, \quad p(L, z, t) = 0 \\ p_z(x, 0, t) = 2\rho u_{tt}, \quad p_z(x, H, t) = 0 \end{cases} \tag{9.8}$$

where $x \in [0, L]$, $z \in [0, H]$ and $t \in [0, \infty)$. In order to solve the above, we let $p = \tilde{p} + \hat{p}$, where $\tilde{p}$ solves

$$\begin{cases} \triangle \tilde{p} = 0 \\ \tilde{p}_x(0, z) = 2\rho s_{tt}, \ \tilde{p}(L, z) = 0 \\ \tilde{p}_z(x, 0) = 0, \ \tilde{p}_z(x, H) = 0 \end{cases} \tag{9.9}$$

and $\hat{p}$ solves

$$\begin{cases} \triangle \hat{p} = 0 \\ \hat{p}_x(0, z) = 0, \ \hat{p}(L, z) = 0 \\ \hat{p}_z(x, 0) = 2\rho u_{tt}, \ \hat{p}_z(x, H) = 0 \end{cases} \tag{9.10}$$

The solution to (9.9) is given by

$$\tilde{p}(x, z, t) = 2\rho s_{tt}(t)(x - L) \tag{9.11}$$

We will use separation of variables to solve (9.10). Thus, we consider a solution of the form $\hat{p}(x,z) = X(x)Z(z)$, which gives us the following ordinary differential equations:

$$\begin{cases} X''(x) &= \mu X(x) \\ -Z''(z) &= \mu Z(z) \end{cases}$$

with boundary conditions

$$\begin{array}{ll} X'(0)Z(z) = 0, & X(L)Z(z) = 0 \\ X(x)Z'(0) = 2\rho u_{tt}, & X(x)Z'(H) = 0 \end{array}$$

Note that the homogeneous case (which gives rise to the Sturm-Louiville problem) is

$$\begin{cases} X''(x) = \mu X(x) \\ X'(0) = X(L) = 0 \end{cases}$$

There are three cases in μ:

1. $\mu = 0$:

   We have $X''(x) = 0 \Rightarrow X(x) = Ax + B$. The boundary conditions give $X'(x) = A = 0$ and $X(L) = B = 0$, and thus we have only the trivial solution.

2. $\mu > 0$:

   We have $X''(x) = \beta^2 X(x)$, $\beta \neq 0$. A general solution is

   $$X(x) = A\cosh\beta x + B\sinh\beta(L-x).$$

   The boundary conditions give

   $$\begin{aligned} X'(0) = -B\beta\cosh(\beta L) = 0 &\Rightarrow B = 0 \\ X(L) = A\cosh\beta L = 0 &\Rightarrow A = 0 \end{aligned}$$

   and thus we have only the trivial solution.

3. $\mu < 0$:

   We have $X''(x) = -\beta^2 X(x)$, $\beta \neq 0$. A general solution is

$$X(x) = A\cos\beta x + B\sin\beta x.$$

   The boundary conditions give

$$X'(0) = B\beta = 0 \Rightarrow B = 0$$
$$X(L) = A\cos\beta L = 0 \Rightarrow \cos\beta L = 0$$

   Thus, we have the family $\beta_n = (n - 1/2)\pi/L \Rightarrow \mu_n = -((n-1/2)\pi/L)^2$ which gives the solutions

$$X_n(x) = \cos(\beta_n x).$$

Now we solve $-Z''(z) = \mu_n Z(z)$, or equivalently

$$\begin{cases} Z''(z) = \beta_n^2 Z(z) \\ Z'(H) = 0 \end{cases}$$

The general solution is given by

$$Z_n(x) = A_n \cosh\beta_n z + B_n \cosh\beta_n(H - z)$$

The boundary condition gives

$$Z_n'(H) = A_n\beta_n \sinh\beta_n H = 0 \Rightarrow A_n = 0$$

Thus,

$$Z_n(x) = B_n \cosh\beta_n(H - z).$$

This gives

$$\hat{p}(x,z) = \sum_{n=1}^{\infty} B_n \cosh\beta_n(H-z)\cos\beta_n x \tag{9.12}$$

We have

$$\hat{p}_z(x,z) = \sum_{n=1}^{\infty} -B_n\beta_n \sinh\beta_n(H-z)\cos\beta_n x$$

$$\hat{p}_z(x,0) = \sum_{n=1}^{\infty} -B_n\beta_n \sinh\beta_n H\cos\beta_n x = 2\rho u_{tt}. \tag{9.13}$$

We also have the fourier series for $2\rho u_{tt}$ given by

$$2\rho u_{tt} = \sum_{n=1}^{\infty} f_n \cos\beta_n x \tag{9.14}$$

where

$$f_n = \frac{2}{L}\int_0^L 2\rho u_{tt}\cos\beta_n x\ dx.$$

By (9.13), we also have

$$f_n = -B_n\beta_n \sinh\beta_n H$$

Solving for B_n, we arrive at

$$B_n = -2\rho\left(\frac{1}{\beta_n\sinh\beta_n H}\right)\left(\frac{2}{L}\int_0^L u_{tt}\cos\beta_n x\ dx\right)$$

Finally, combining the above equation with (9.11) and (9.12), we arrive at

$$p(x,z,t) = \tilde{p}+\hat{p} = 2\rho s_{tt}(t)(x-L) + \sum_{n=1}^{\infty} B_n\cosh\beta_n(H-z)\cos\beta_n x. \tag{9.15}$$

If we denote the orthonormal family $\phi_n(x) = \sqrt{\frac{2}{L}}\cos\beta_n x$, we can rewrite (9.15) at $z=0$ as

$$p(x,0,t) = 2\rho s_{tt}(t)(x-L) - 2\rho M_f\{u_{tt}(x,t)\} \tag{9.16}$$

where

$$M_f\{u_{tt}(x,t)\} = \sum_{n=1}^{\infty} Q_2(\beta_n) <\ u_{tt}(x,t), \phi_n(x)\ >_{L_2(0,L)} \phi_n(x) \tag{9.17}$$

$$Q_2(\beta_n) = \frac{\coth \beta_n H}{\beta_n}$$

The subscript on Q denotes the dimension of the fluid. It's significance will be seen in the last section of this chapter.

9.2 Simple BM model

For the purposes of this chapter, we will start with a simple mechanical model of the cochlear partition as just the basilar membrane and ignore the effects of the tectorial membrane (TM) and Organ of Corti (OC). This is simply to at first illustrate the analysis and dynamics of coupling the fluid with the tissue. Thus, we will consider a passive ear where the micromechanics of the OC *do not* affect the macromechanical vibration of the basilar membrane. Also, the TM (the "second filter") will be ignored.

Helmholtz [15] envisioned the BM like strings on a piano, where each string vibrates independently of the others. This is more in line with the terminology of the basilar *membrane*. However, studies from [14] have shown that the restoring force upon deflection of the BM behaves more like that of a "plate" or "beam," where there is no resting tension on the BM. Thus, a more accurate description would be basilar *plate*. Due to the simplicity of Helmholtz's description, however, we will model the BM as a collection of mass/spring systems.

Is there longitudinal coupling along the BM? Studies from Békésy [36] have shown it can occur either through mechanical or fluid coupling, unlike Helmholtz's piano string description where each string vibrated independently of the others. Voldrich [35] was able to demonstrate that the coupling is dominated by the fluid (See Voldrich), and thus we ignore the mechanical

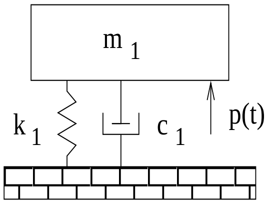

Figure 9.3: Cross section micromechanics of the basilar membrane (m_1).

coupling along the BM. We have the following description of the BM:

$$\begin{cases} m(x)u_{tt} + c(x)u_t + k(x)u = p(x,0,t), \ x \in [0,L] \\ u(x,0) = u_t(x,0) = 0 \end{cases} \tag{9.18}$$

where $p(x,z,t)$ is the pressure from the fluid and $u(x,t)$ is BM displacement.

9.2.1 Steady State Formulation

In this paper, we will use the terms steady state, standing wave and planar wave interchangeably. Thus, consider steady state solutions of the form

$$\begin{aligned} p(x,z,t) &= \mathrm{Im}\{P(x,z)e^{i\omega t}\} \\ u(x,t) &= \mathrm{Im}\{U(x)e^{i\omega t}\} \\ s(t) &= \mathrm{Im}\{Se^{i\omega t}\} \end{aligned}$$

Plugging this into the fluid equation (9.8), we have

$$\begin{cases} \triangle P = 0 \\ P_x(0,z) = 2\rho(i\omega)^2 S, \ P(L,z) = 0 \\ P_z(x,0) = 2\rho(i\omega)^2 U(x), \ P_z(x,H) = 0 \end{cases} \tag{9.19}$$

Plugging the above into the equations (9.18), we have

$$\begin{aligned} P(x,0) &= [-m(x)\omega^2 + i\omega c(x) + k(x)]U(x) \\ &= [i\omega m(x) + c(x) + k(x)/i\omega]i\omega U(x) \\ &= Z_{CP}(x)i\omega U(x) \end{aligned} \tag{9.20}$$

where

$$Z_{CP}(x) = i\omega m(x) + c(x) + k(x)/i\omega$$

is called the *impedance* for the cochlear partition. If you solve for $Z_{CP}(x)$ in (9.20), you get

$$Z_{CP}(x) = \frac{P(x,0)}{i\omega U(x)}$$

Thus, the impedance is the ratio of pressure over velocity and is in the units of damping. A high impedance means a lot of pressure but little movement and a low impedance is a little pressure for a lot of movement. Different tissue models will give different impedances. It is a complex function and contains damping as well as *phase information.*

Going back to the steady state equation (9.19), we substitute for $U(x)$ giving

$$\begin{cases} \triangle P = 0 \\ P_x(0,z) = 2\rho(i\omega)^2 S, \ P(L,z) = 0 \\ P_z(x,0) = \frac{2\rho i\omega}{Z_{CP}(x)} P(x,0), \ P_z(x,H) = 0 \end{cases} \tag{9.21}$$

A plot of a steady state solution for a single 4 kHz tone at 50 dB SPL is shown in Figure 9.4.

9.2.2 Time Domain Formulation

Combining the fluid (9.16) and tissue (9.18), we arrive finally at

$$\begin{cases} (m(x) + 2\rho M_f)\, u_{tt} + c(x)u_t + k(x)u = 2\rho s_{tt}(t)(x-L) \\ u(x,0) = u_t(x,0) = 0 \end{cases} \tag{9.22}$$

where M_f is given in (9.17).

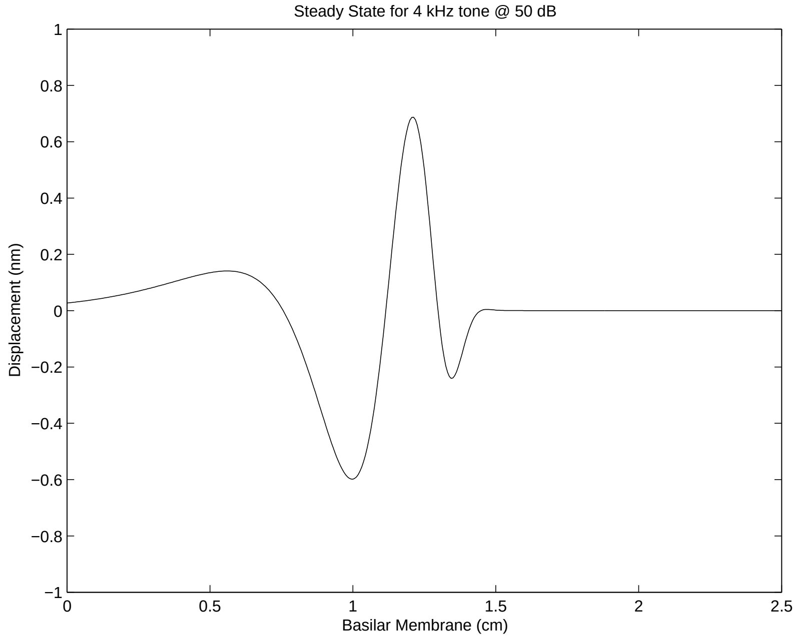

Figure 9.4: Steady state BM response to a 4 kHz tone with intensity of 50 dB SPL.

9.2.3 Natural Boundary Condition

What are the natural boundary conditions at $x = 0$ and $x = L$? Let's look at the basal end first. We have

$$p_x(0, z, t) = 2\rho s_{tt} \Longrightarrow p_{xz}(0, 0, t) = 0$$

as well as

$$p_z(x, 0, t) = 2\rho u_{tt}(x, t) \Rightarrow p_{zx}(0, 0, t) = 2\rho u_{ttx}(0, t)$$

Applying equality of mixed partials and the initial conditions, we get

$$u_{ttx}(0, t) = 0 \Longrightarrow u_x(0, t) = 0.$$

For the apical end, we have

$$p(L, z, t) = 0 \Longrightarrow p(L, 0, t) = 0$$

Plugging this into the BM boundary condition, we get

$$m(L)u_{tt}(L, t) + c(L)u_t(L, t) + k(L)u(L, t) = 0$$

along with the zero initial conditions in (9.18). Thus, we easily see that the only solution is $u(L, t) = 0$. So, the full BM equation is given by

$$\begin{cases} (m(x) + 2\rho M_f)\, u_{tt} + c(x)u_t + k(x)u = 2\rho s_{tt}(t)(x - L), \; x \in [0, L] \\ u_x(0, t) = 0, \; u(L, t) = 0 \\ u(x, 0) = u_t(x, 0) = 0 \end{cases}$$

Remark 9.2.1. The natural boundary conditions can also be understood in relation to the fourier series of u_{tt} in (9.14). The basis functions $\phi^k(x) = \cos \beta_k x$ all satisfy $\phi^k(L) = 0$ and $\phi^k_x(0) = 0$. We can, and will, omit them most of the time. They become particularly important in studying the Neumann apical boundary condition (see [38]).

9.2.4 Dispersion Relation

For analysis of the dispersion relation, we extend the x-domain in (9.8) to all of $\mathbb{R}$, giving

$$\begin{cases} \triangle p = 0, \; x \in \mathbb{R}, \; z \in [0, H], \; t \in [0, \infty) \\ p_z(x, 0) = 2\rho u_{tt}, \; p_z(x, H) = 0 \end{cases}$$

with BM motion given by

$$p(x, 0, t) = m u_{tt} + c u_t + k u$$

where m, c and k are now constants. Begin with solutions of the form

$$p(x, z, t) = p_0(z)e^{i(\kappa x - \omega t)}, \; u(x, t) = u_0 e^{i(\kappa x - \omega t)}$$

where κ is the wavenumber for x. We have

$$\begin{aligned} \triangle p = 0 &\implies -\kappa^2 p_0(z) e^{i(\kappa x - \omega t)} + p_0''(z) e^{i(\kappa x - \omega t)} = 0 \\ &\implies p_0''(z) - \kappa^2 p_0(z) = 0 \end{aligned}$$

The boundary conditions give

$$\begin{aligned} p_z(x,0,t) = 2\rho u_{tt} &\implies p_0'(0) = -2\rho u_0 \omega^2 \\ p_z(x,H,t) = 0 &\implies p_0'(H) = 0 \end{aligned}$$

Thus, we solve the following second-order boundary value problem

$$\begin{cases} p_0''(z) - \kappa^2 p_0(z) = 0 \\ p_0'(0) = -2\rho u_0 \omega^2 \\ p_0'(H) = 0 \end{cases}$$

Our solution to the above is

$$\begin{aligned} p_0(z) &= a e^{\kappa z} + b e^{-\kappa z} \\ a &= \left(\frac{-2\rho u_0 \omega^2}{\kappa}\right)\left(\frac{1}{1 - e^{2\kappa H}}\right) \\ b &= \left(\frac{-2\rho u_0 \omega^2}{\kappa}\right)\left(\frac{e^{2\kappa H}}{1 - e^{2\kappa H}}\right) \end{aligned}$$

This gives

$$\begin{aligned} p_0(0) &= a + b \\ &= \left(\frac{-2\rho u_0 \omega^2}{\kappa}\right)\left(\frac{1 + e^{2\kappa H}}{1 - e^{2\kappa H}}\right) \\ &= \left(\frac{-2\rho u_0 \omega^2}{\kappa}\right)\coth(-\kappa H) \\ &= 2\rho u_0 \omega^2 Q_2(\kappa) \end{aligned}$$

Now, plugging the wave solutions into the BM boundary condition gives

$$p_0(0) = (-m\omega^2 - ic\omega + k)u_0$$

which leads to the following dispersion relation

$$(m + 2\rho Q_2(\kappa))\omega^2 + ic\omega - k = 0$$

or

$$\omega = \frac{-ic \pm \sqrt{4k(m + 2\rho Q_2(\kappa)) - c^2}}{2(m + 2\rho Q_2(\kappa))}$$

Since $4k(m+2\rho Q_2(\kappa)) > 0$, $\Im\{\omega\} < 0$, and thus the system is *dissipative.* We have $\lim_{\kappa\to 0^+} Q_2(\kappa) = +\infty$, and thus

$$\lim_{\kappa\to 0^+} \Im\{w\} = 0$$

which shows the slow decay of long waves. If we set the damping coefficient c to zero, we have

$$\omega(\kappa) = \pm\sqrt{\frac{k}{m + 2\rho Q_2(\kappa)}}$$

which is what is to be expected. The group velocity is given by

$$\omega'(\kappa) = \pm\frac{\rho\sqrt{k}[H\operatorname{csch}^2\kappa H + Q_2(\kappa)]}{\kappa(m + 2\rho Q_2(\kappa))^{3/2}}$$

Note that $\omega'(\kappa) = O(\kappa^{-3/2})$ as $\kappa \to \infty$, so *short waves do not disperse as fast as long waves.* How fast do the long waves disperse? We have

$$\begin{aligned}
\omega'(0) &= \lim_{\kappa\to 0^+} \frac{\omega(\kappa)}{\kappa} \\
&= \lim_{\kappa\to 0^+} \pm\frac{1}{\kappa}\left(\frac{k}{m + 2\rho Q_2(\kappa)}\right)^{1/2} \\
&= \lim_{\kappa\to 0^+} \pm\left(\frac{k}{\kappa^2 m + 2\rho\kappa^2 Q_2(\kappa)}\right)^{1/2} \\
&= \pm\sqrt{\frac{kH}{2\rho}}
\end{aligned}$$

where we used the fact that

$$\lim_{\kappa\to 0^+} \kappa^2 Q_2(\kappa) = \frac{1}{H}.$$

This is an important limit that will show up often in our analysis of the problem. How fast do the long waves decay? From this limit, we can show

$$\Im\{\omega\} = O(\kappa^2)$$

as $\kappa \to 0^+$.

9.3 The TM/BM Model

We will now add more structure to the cochlear partition by adding the tectorial membrane. As was mentioned in the previous chapter, it is postulated that the tectorial membrane adds a second filter to the mechanical vibration of the cochlear partition. The basilar membrane and tectorial membrane are coupled through the outer hair cells. Thus, our construction follows the form of Neely and Kim's paper [25] and is shown in Figure 9.5.

Let $\xi(x,t) = (u(x,t), v(x,t))$ be BM and TM displacement, respectively. If we construct the corresponding free-body diagram, we arrive at the cochlear partition equations

$$M_p\xi_{tt} + C_p\xi_t + K_p\xi = F \tag{9.23}$$

where

$$M_p = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad C_p = \begin{bmatrix} c_1 + c_3 & -c_3 \\ -c_3 & c_2 + c_3 \end{bmatrix}, \quad K_p = \begin{bmatrix} k_1 + k_3 & -k_3 \\ -k_3 & k_2 + k_3 \end{bmatrix}$$

and

$$F = \begin{bmatrix} p(x,0,t) \\ 0 \end{bmatrix}$$

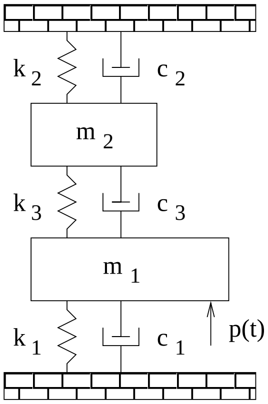

Figure 9.5: Cross section micromechanics of the cochlea. The mass m_1 represents a cross section of the BM, while mass m_2 is a cross section of the TM. (Reconstructed from Figure 3 in Neely and Kim [25])

9.3.1 Steady State Formulation

Revisiting equation (9.21), we see that the only modification needed to compute the steady state solution of the TM/BM model is to construct the cochlear partition impedance Z_{CP}. To this end, we substitute solutions of the form

$$p(t) = Pe^{i\omega t}, \quad u(x,t) = U(x)e^{i\omega t}, \quad v(x,t) = V(x)e^{i\omega t}$$

into the mechanical equations (9.23) and solve for $Z_{CP} = P/i\omega U$. You will arrive at

$$\begin{aligned} Z_{CP} &= Z_1 + Z_2 Z_3/(Z_2 + Z_3) \\ Z_1 &= k_1/i\omega + c_1 + i\omega m_1 \\ Z_2 &= k_2/i\omega + c_2 + i\omega m_2 \\ Z_3 &= k_3/i\omega + c_3 \end{aligned} \tag{9.24}$$

This, in conjunction with equation (9.21), gives the steady state solution for the TM/BM model.

9.3.2 Time Domain Formulation

Combining (9.16) with (9.23), we have

$$\begin{cases} M_p^f \xi_{tt} + C_p \xi_t + K_p \xi = F \\ \xi(x,0) = \xi_t(x,0) = 0 \end{cases} \tag{9.25}$$

where

$$M_p^f = \begin{bmatrix} m_1 + 2\rho M_f & 0 \\ 0 & m_2 \end{bmatrix}$$

and M_f is given in (9.17).

9.3.3 Dispersion Relation

To construct the dispersion relation for the TM/BM model, we will follow a similar analysis to that in section 9.2.4. We have

$$\frac{p_0(0)}{i\omega u_0} = Z_{CP} \Longrightarrow 2\rho\omega^2 Q_2(\kappa) = i\omega Z_{CP}$$

where now Z_{CP} is given by equation (9.24). Thus, letting $\mathcal{Z}_i = i\omega Z_i$, we have

$$2\rho\omega^2 Q_2(\kappa)(\mathcal{Z}_2 + \mathcal{Z}_3) = \mathcal{Z}_1(\mathcal{Z}_2 + \mathcal{Z}_3) + \mathcal{Z}_2\mathcal{Z}_3.$$

We care again about the long-wave behaviour. Multiplying both sides by κ^2 and taking a limit as $\kappa \to 0^+$, we arrive at

$$\omega^2(0)(\mathcal{Z}_2(0) + \mathcal{Z}_3(0)) = 0.$$

Thus, we have a double zero at $\omega_1(0) = 0$ and

$$\omega_2(0) = \frac{i(c_2 + c_3) \pm \sqrt{4m_2(k_2 + k_3) - (c_2 + c_3)^2}}{2m_2}$$

Note that $\Im\{\omega_1\} = 0$ means that we still have slow decay of long waves. If we consider the undamped case ($c_i = 0$), then

$$\omega_2(0) = \pm\sqrt{\frac{k_2 + k_3}{m_2}}$$

So how fast do long waves disperse? We have 4 waves: 2 the BM waves, 2 the TM waves. Let us look at the BM waves $\omega_1(\kappa) = \omega_1'(0)\kappa$ for $\kappa \ll 1$. We have

$$2\rho\omega_1^2(\kappa)Q_2(\kappa) = 2\rho\omega_1'(0)^2\kappa^2 Q_2(\kappa) \to (2\rho/H)\omega_1'(0)^2$$

$$\begin{aligned}
\mathcal{Z}_i &= -m_i\omega_1^2(\kappa) + k_i \\
&= -m_i\kappa^2\omega_1'(0)^2 + k_i \\
&\to k_i
\end{aligned}$$

Thus, we have

$$(2\rho/H)\omega_1'(0)^2(k_2+k_3) = k_1(k_2+k_3) + k_2k_3$$

$$\omega_1'(0) = \pm\sqrt{\frac{H\left(k_1 + \frac{k_2k_3}{k_2+k_3}\right)}{2\rho}}$$

Now let us look at the TM waves $\omega_2(\kappa)$. It is easiest in this case to consider the equation above for $\kappa \ll 1$, giving

$$\begin{aligned}
\omega^2(\mathcal{Z}_2 + \mathcal{Z}_3) = 0 &\implies \omega^2[-m_2\omega^2 + k_2 + k_3] = 0 \\
&\implies [-m_2\omega^4 + \omega^2(k_2+k_3)]' = 0
\end{aligned}$$

Thus, as $\kappa \to 0^+$, we have

$$2\omega_2(0)\omega_2'(0)[k_2 + k_3 - 2m_2\omega_2^2(0)] = 0$$

Now, since $k_2 + k_3 - 2m_2\omega_2^2(0) \neq 0$ and $\omega_2(0) \neq 0$,

$$\omega_2'(0) = 0.$$

Let us look at the second derivative to get more information. Denoting $\omega_2(0) = \omega_{2,0}$, we have

$$\omega_2(\kappa) = \omega_{2,0} + \frac{\kappa^2}{2}\omega_{2,0}'$$

Substituting above, we have

$$2\rho\omega_2^2(\kappa)Q_2(\kappa) = 2\rho Q_2(\kappa)[\omega_{2,0}^2 + O(\kappa^2) + \text{h.o.t.}]$$

$$\begin{aligned} \mathcal{Z}_2 + \mathcal{Z}_3 &= -m_2\omega_{2,0}^2 - m_2\omega_{2,0}\omega_{2,0}''\kappa^2 + \text{h.o.t.} + k_2 + k_3 \\ &= -m_2\omega_{2,0}\omega_{2,0}''\kappa^2 + \text{h.o.t.} \end{aligned}$$

by definition of $\omega_{2,0}$. Thus, as $\kappa \to 0^+$, the left-hand side converges to

$$-\frac{2\rho}{H}m_2\omega_{2,0}^3\omega_{2,0}'' = -\frac{2\rho}{H}(k_2 + k_3)\omega_{2,0}\omega_{2,0}''$$

For the right-hand side, we have

$$\mathcal{Z}_1(\mathcal{Z}_2 + \mathcal{Z}_3) \to 0$$

$$\begin{aligned} \mathcal{Z}_2\mathcal{Z}_3 &= (-m_2\omega_{2,0}^2 + k_2 + \text{h.o.t.})k_3 \\ &= (-k_3 + \text{h.o.t.})k_3 \\ &\to -k_3^2 \end{aligned}$$

Finally, we have

$$\omega_2''(0) = \left(\frac{H}{2\rho}\right)\sqrt{\frac{m_2}{k_2 + k_3}}\left[\frac{k_3^2}{k_2 + k_3}\right]$$

9.4 Long-Wave Approximation (One-Dimensional Model)

For completeness and an overall understanding of the problem, we will show how what is known as the *long-wave* model is derived. We start with a similar simplification as was done in the previous section, namely we extend the z-domain to all of $\mathbb{R}$, giving

$$\begin{cases} \triangle p = 0, \ x \in [0, L], \ z \in \mathbb{R}, \ t \in [0, \infty) \\ p_x(0, z) = 2\rho(i\omega)^2 Se^{i(lz - \omega t)}, \ p(L, z) = 0 \end{cases}$$

with BM motion given by

$$p(x,0,t) = mu_{tt} + cu_t + ku$$

where m, c and k are now constants. We begin again with solutions of the form

$$p(x,z,t) = p_0(x)e^{i(lz-\omega t)}.$$

where now l is the wavenumber for z. Thus, we have

$$\begin{cases} p_0''(x) - l^2 p_0(x) = 0 \\ p_0'(0) = 2\rho(i\omega)^2 S, \;\; p_0(L) = 0 \end{cases}$$

How do the wavenumbers κ and l relate to each other? If we assume periodicity along the x and z and plug $p(x,z,t) = p_0 e^{i(\kappa x + lz - \omega t)}$ into the Laplacian, we arrive at

$$\kappa^2 + l^2 = 0$$

and thus we have

$$p_0''(x) + \kappa^2 p_0(x) = 0.$$

Now, *this is the key step in reducing the problem to one-dimension.* We want to look at long waves, i.e. $\kappa \ll 1$. Thus, from the dispersion relation above, we have

$$\begin{aligned} \frac{p_0(0)}{i\omega u_0} = Z_{BM} \quad &\Longrightarrow \quad -2\rho i\omega Q_2(\kappa) = Z_{BM} \\ &\Longrightarrow \quad \kappa^2 = \left(\frac{-2\rho i\omega}{Z_{BM}}\right)\kappa^2 Q_2(\kappa) \\ &\Longrightarrow \quad \kappa^2 \approx \frac{-2\rho i\omega}{HZ_{BM}}, \text{ for } \kappa \ll 1 \end{aligned}$$

Substituting this above, we finally arrive at the *one-dimensional long-wave approximation* (dropping the subscript on p)

$$\begin{cases} p''(x) - \left(\frac{2\rho i\omega}{HZ_{BM}}\right) p(x) = 0 \\ p'(0) = 2\rho(i\omega)^2 S, \;\; p(L) = 0 \end{cases}$$

This is the steady-state case of the model, and it is easily shown that the time-domain analogue is given by

$$\begin{cases} p_{xx} - \frac{2\rho}{H} u_{tt} = 0 \\ p_x(0,t) = 2\rho s_{tt}, \ p(L,t) = 0 \\ u(x,0) = u_t(x,0) = 0 \end{cases} \tag{9.26}$$

The time domain model (9.26) was studied by Xin in [39].

Remark 9.4.1. For the two-dimensional model, we defined

$$Q_2(\kappa) = \frac{\coth \kappa H}{\kappa}.$$

For $\kappa \ll 1$, $Q_2(\kappa) \approx 1/H\kappa^2$, so we define

$$Q_1(\kappa) = \frac{1}{H\kappa^2}$$

Given this, we have

$$\lim_{\kappa \to 0^+} \kappa^2 Q_1(\kappa) = \frac{1}{H}$$

which is the same for $Q_2(\kappa)$. Thus, *long waves disperse/decay at the same rate as the two-dimensional model.* For more discussion on the significance of Q_1 versus Q_2, see de Boer [7].

Chapter 10

The Middle Ear

During much of the earlier research into the cochlea, the models created were all steady state models [24, 25]. The problem comes in moving from the linear ear to the nonlinear ear. Once we are out of the wonderfully peaceful and comfortable place of the linear regime, with all of its power and tools, we are forced to look at time domain solutions. It is interesting that in this problem, immediately, even in the simplest of cases of sinusoidal input on one tone to the ear, we run into a very serious issue: that of a dispersive instability at the apex of the cochlea.

It is instructive to cover this topic in more detail in the chapter on the middle ear since, at least for sinusoidal inputs, our time domain formulation of the middle ear reduces the dispersive instability. The time domain formulation of the middle ear [13] is given as

$$m_e s_{tt} + c_e s_t + k_e s = p_e(t) \tag{10.1}$$

where $s(t)$ is stapes displacement and $p_e(t)$ is pressure at the eardrum. As was mentioned in the previous chapter, this is coupled to the BM by

$$p_x(0, z, t) = 2\rho s_{tt}$$

where $p(x, z, t)$ is fluid pressure in the cochlea and ρ is fluid density. This is a linear middle ear and as such can be formulated in the frequency domain.

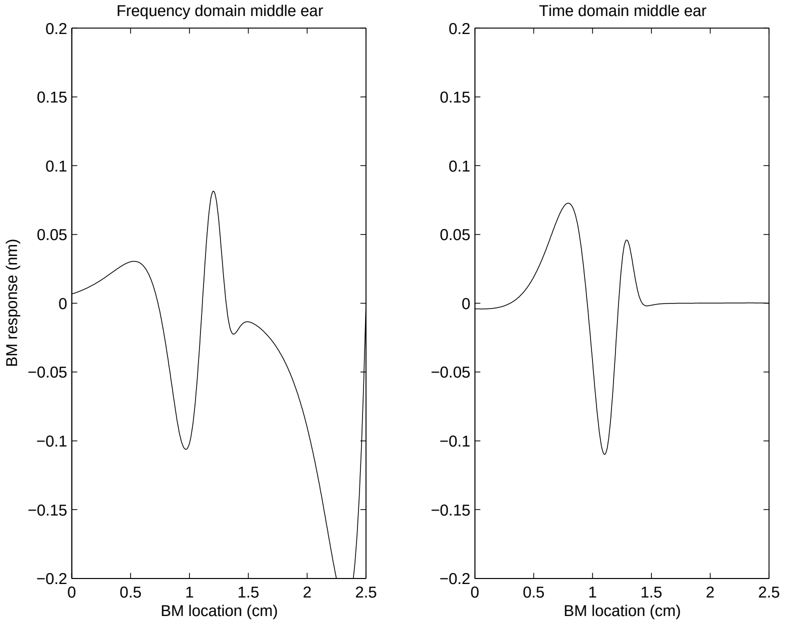

Figure 10.1: Both plots are BM response time domain snapshots at $t = 20$ ms. The left snapshot is a frequency domain middle ear, while the right is a time domain middle ear.

To illustrate the difference between using a time domain middle ear with a time domain BM model or a frequency domain middle ear with a time domain model, consider Figure 10.1. Both plots solve the time domain system (9.22) with a 4 kHz tone at 50 dB SPL. The plot on the left is with a frequency domain middle ear and the plot on the right is a time domain middle ear. Notice the large dispersive tail using the frequency domain middle ear.

In section 1, we cover the filtering characteristics of our middle ear model (10.1) to show how it amplifies the signal. In section 2, we derive the filter for the basilar membrane and then compare the time domain response with the steady state response without the middle ear. Finally, in section 3, we combine the middle ear with the basilar membrane, derive its filter, and then perform the same comparison between time domain and steady state to show the middle ear's ability to reduce the dispersive instability.

10.1 Vibration Analysis: Middle Ear

Considering equation (10.1), we assume harmonic inputs $s(t) = Se^{i\omega t}$ and $p_e(t) = Pe^{i\omega t}$. Substituting these into (10.1) and dividing by $e^{i\omega t}$, we get

$$\begin{aligned} & (m_e(i\omega)^2 + i\omega c_e + k_e)S = P \\ \Longrightarrow \quad & \tfrac{S}{P} = \tfrac{1}{m_e(i\omega)^2 + i\omega c_e + k_e} \end{aligned} \tag{10.2}$$

The absolute value of this expression gives

$$\left|\frac{S}{P}\right| = \frac{1/k_e}{\sqrt{(1-r_e^2)^2 + (2\zeta_e r_e)^2}} \equiv \frac{1}{k_e}\phi_e(r_e) \tag{10.3}$$

where the dimensionless parameters $r_e,\ \zeta_e$ are given by

$$\begin{aligned} r_e = \quad & \tfrac{\omega}{\omega_e} \quad && = \text{normalized frequency ratio} \\ \zeta_e = \quad & \tfrac{c_e}{2m_e\omega_e} \quad && = \text{normalized damping ratio} \end{aligned}$$

and $\omega_e = \sqrt{k_e/m_e}$ is the natural frequency of the middle ear (taken to be $2\pi f_e$, where $f_e = 4\ kHz$). The damping ratio for the ear is given as $\zeta_e = 0.7$. The natural frequency and damping ratio for the middle ear were taken from [13] (see Table A.2). Now, we want to consider the following ratio

$$\frac{\text{Force at stapes}}{\text{Force at eardrum}}$$

The force at the stapes is given by $2\rho s_{tt}$, or $2\rho\omega^2 S$ in the frequency domain. Thus, we have

$$\left|\frac{2\rho\omega^2 S}{P}\right| = \frac{2\rho}{k_e}\omega^2\phi_e = \frac{2\rho}{k_e}\omega_e^2 r_e^2\phi_e = \frac{2\rho}{m_e}r_e^2\phi_e \tag{10.4}$$

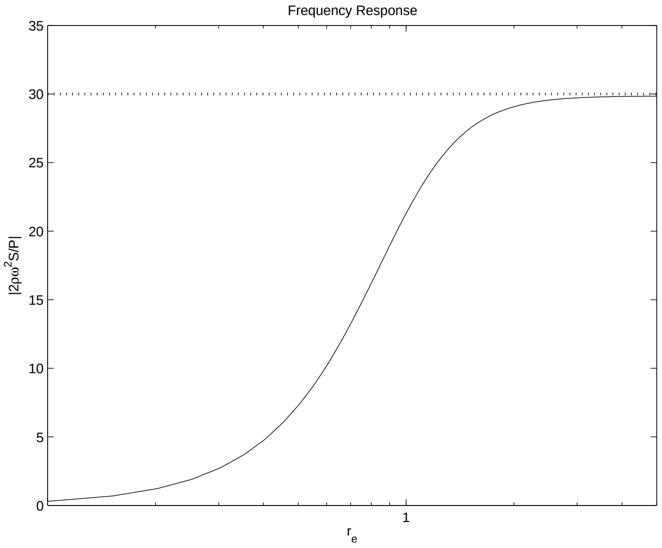

Figure 10.2: This is a plot of equation (10.4) for the middle ear with damping ratio $\zeta_e = 0.7$. This shows the middle ear is a vibration amplifier over most frequencies.

This is the middle ear's frequency response $|H_e(i\omega)|$ and gives the filtering characteristics of the middle ear. The amplification factor $2\rho/m_e$ is a ratio of the size and mass of the eardrum to the stapes. Using the parameters from Table A.2, the amplification factor becomes

$$\frac{2\rho}{m_e} = 30$$

and thus the middle ear *amplifies* the signal. See Figure 10.2 for a plot of the middle ear frequency response.

10.2 Vibration Analysis: Basilar Membrane

Consider the simple BM model from equation (9.22). We will ignore the fluid for simplicity and just look at one slice of the BM at location x, giving the spring-mass-damping system

$$\begin{cases} m_b u_{tt} + c_b u_t + k_b u = f(t) \\ u(0) = u_0, \;\; u_t(0) = u_1 \end{cases} \tag{10.5}$$

where $f(t) = 2\rho s_{tt}(t)(x - L)$.

10.2.1 Steady State

As in the previous section, we assume solutions of the form

$$u(t) = Ue^{i\omega t}, \quad f(t) = Fe^{i\omega t}$$

Following a similar analysis as above, we arrive at

$$\left|\frac{U}{F}\right| = \frac{1/k_b}{\sqrt{(1 - r_b^2)^2 + (2\zeta_b r_b)^2}} \tag{10.6}$$

which is the frequency response $|H_b(i\omega)|$ for the basilar membrane. It is a little different in form from (10.4) in that it is a ratio of displacement to force. It is instructive to plot what is known as the *normalized displacement* by multiplying (10.6) by k_b, giving

$$\left|\frac{Uk_b}{F}\right| = \frac{1}{\sqrt{(1 - r_b^2)^2 + (2\zeta_b r_b)^2}} \equiv \phi_b(r_b) \tag{10.7}$$

The normalized displacement $\phi_b(r_b)$ is now fully nondimensional and order unity. It is the ratio of the input displacement F/k_b (as if there were no mass or damping) to the output displacement U. See Figure 10.3 for an illustration.

Figure 10.4 shows plots of the normalized displacement at various points along the BM using parameters from Table A.2. These curves are particularly useful for tuning the system.

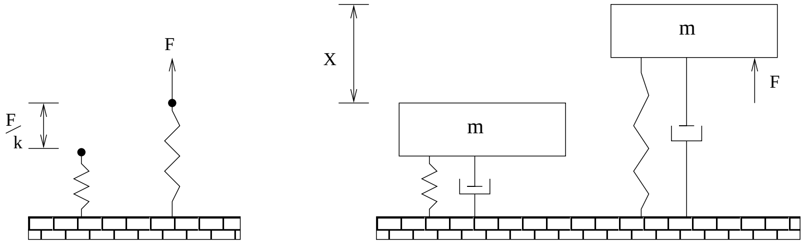

Figure 10.3: The left plot is a force F acting directly on a spring, showing its deflection F/k. The right plot shows the maximum steady state response X of a spring-mass-damper system under the same force F. The ratio Xk/F is called the normalized displacement and gives the signal amplification/suppression characteristics of the system.

10.2.2 Time Domain

The previous section discussed the steady state response of the system given in (10.5). In any design of a vibrational system, however, one must not ignore transient effects. The following quote from Inman [16] (page 69) stresses the importance of this issue:

> Usually, devices are designed and analyzed based on the steady-state response, but the transient should always be checked to make sure that it is reasonable to ignore it, or if it should, in fact, be considered seriously.

It should be noted as well that near the apex of the cochlea, the transient part of the solution is *much* larger than the steady state part. Since the damping is very small there, the larger transient part takes a long time to decay. Thus, we must, as Inman stated, look a little closer at how transients effect the basilar membrane.

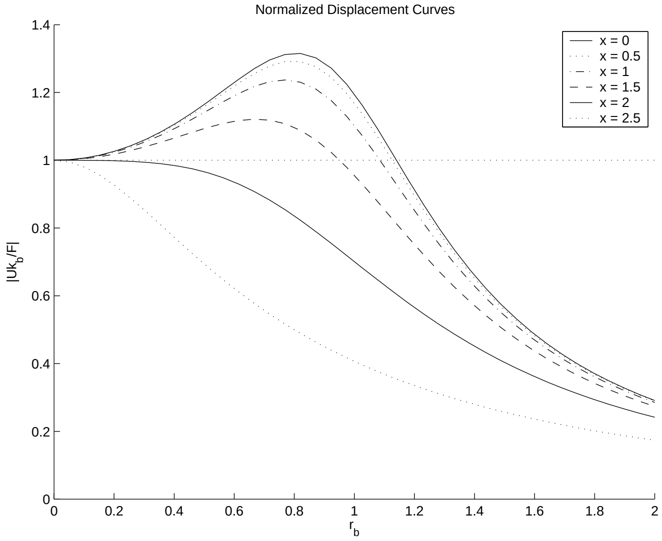

Figure 10.4: Plots of normalized displacement curves for various locations along the basilar membrane.

10.2.2.1 Sinusoidal Inputs

For illustrative purposes, we will simplify the problem to the undamped case. Considering sinusoidal inputs, we have

$$m_b u_{tt} + k_b u = F\sin(\omega t + \psi) \tag{10.8}$$

We allow a phase shift to make the problem more general (It turns out the phase shift is very important here). The general solution is given by

$$u(t) = U_T\sin(\omega_b t + \phi) + U_F\sin(\omega t + \psi) \tag{10.9}$$

where U_T is the homogeneous (transient) part and U_F is the particular (steady state) part. The use of transient and steady state comes from the fact that if damping were present, there would be an $e^{-\zeta_b t}$ scaling the U_T term. The normalized displacement is given by

$$\left|\frac{U_F k_b}{F}\right| = \frac{1}{|1 - r_b^2|}$$

Now, we want to compare the relative sizes of U_T to U_F. Continuing, we choose the initial conditions as

$$u(0) = U_F u_0, \quad u_t(0) = U_F \omega u_1$$

Substituting into (10.9), we have

$$\begin{aligned} u(0) &= U_T \sin\phi + U_F \sin\psi = U_F u_0 \\ u_t(0) &= \omega_b U_T \cos\phi + \omega U_F \cos\psi = U_F \omega u_1 \end{aligned}$$

This gives

$$\begin{aligned} U_T \sin\phi &= U_F(u_0 - \sin\psi) \\ U_T \cos\phi &= U_F r_b (u_1 - \cos\psi) \end{aligned}$$

and thus

$$\frac{U_T}{U_F} = \sqrt{(u_0 - \sin\psi)^2 + r_b^2 (u_1 - \cos\psi)^2}$$

Considering now the initial conditions $u_0 = 0$ and $u_1 = 0$, we have

$$\frac{U_T}{U_F} = \sqrt{\sin^2\psi + r_b^2 \cos^2\psi} \tag{10.10}$$

See Figure 10.5 for an illustration of this equation. So we see that the initial phase shift of the forcing function makes a *huge* difference. If $\psi = 0, \pi$, then we have

$$U_T / U_F = r_b \tag{10.11}$$

At the apex, $r_b \gg 1$ for most audible frequencies, so $U_T \gg U_F$. For the actual size of the transient part when $\psi = 0, \pi$, we combine equations (10.11) and

$$\left| \frac{U_F k_b}{F} \right| = \frac{1}{|1 - r_b^2|}$$

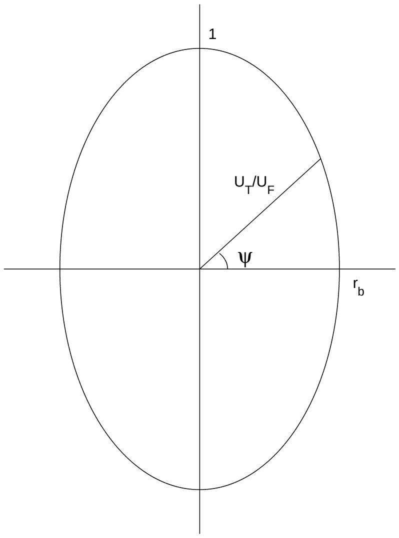

Figure 10.5: Illustration of equation (10.10).

from the previous section, arriving at

$$\left|\frac{U_T k_b}{F}\right| = \frac{r_b}{|1 - r_b^2|}$$

For the time domain response, we reinsert damping into equation (10.8) and consider $\psi = 0$, giving

$$m_b u_{tt} + c_b u_t + k_b u = F sin(\omega t)$$

We solve the problem in the time domain and compute the normalized displacement as

$$\left|\frac{u(t)k_b}{F}\right|_{\text{max}}$$

Note that the time domain response $u(t)$ contains both the transient part U_T and the steady state part U_F. A plot of normalized displacement for the steady state and time domain case is given in the left plot of Figure 10.6 for an apical point on the BM.

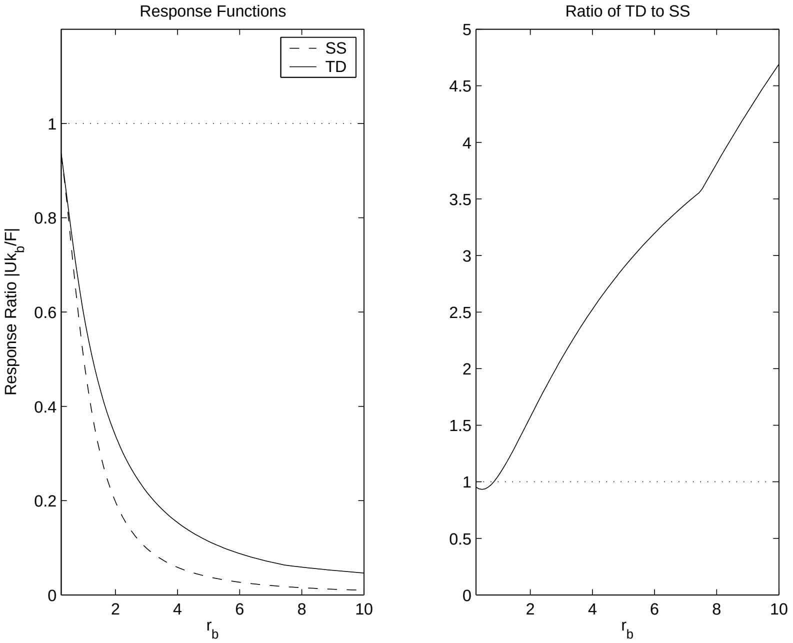

Figure 10.6: The left plot consists of normalized displacement curves for an apical site on the BM. The right plot is the ratio of the maximum time domain response to the steady state response.

10.2.2.2 Time Domain/Steady State Ratio

When $\psi = 0$, we arrived at

$$\frac{U_T}{U_F} = r_b$$

Thus, we get

$$\frac{|u(t)|\text{max}}{U_F} = \frac{U_F + U_T}{U_F} = 1 + \frac{U_T}{U_F} = 1 + r_b$$

This is the case without damping. If we add damping, compute the time domain response and form this ratio, we arrive at the plot on the right in Figure 10.6. We see that the time domain response relative to the steady state response grows nearly linearly as r_b increases, similar to the undamped case.

When the ratio is 1, there are no transient effects in the solution. This gives us a handle on how large the transient effects are relative to the steady state solution. In the next section, we will couple the middle ear with the basilar membrane and repeat this analysis. The result will be a reduction in this ratio, showing the middle ears ability to reduce transients, even though it amplifies the signal.

10.3 Vibration Analysis: Basilar Membrane with Middle Ear

Since the middle ear and cochlea act in series, the frequency response of the combined system is simply the frequency responses of both systems multiplied together. Thus, combining equations (10.6) and (10.4), we have the frequency response for the combined system as

$$\left|\frac{U}{P}\right| \equiv |H_{eb}(i\omega)| = |H_e(i\omega)| \cdot |H_b(i\omega)| = \frac{2\rho}{m_e k_b} r_e^2 \phi_e \phi_b$$

The left plot in Figure 10.7 shows the normalized displacement for the combined system under sinusoidal input. Note that the normalized displacement for the combined system is

$$\left|\frac{U k_b}{P}\right| = k_b |H_{eb}(i\omega)| = \frac{2\rho}{m_e} r_e^2 \phi_e \phi_b$$

The right plot of Figure 10.7 gives the ratio of the maximum time domain response to the steady state response. When compared with Figure 10.6, we see that the time domain/steady state ratio is reduced for relatively large frequencies. *This shows the middle ear acting as a transient absorber.*

It should be noted that this was a study of transient behaviour for a particular input, namely sinusoidal. A more thorough treatment of other transient phenomenon, such as clicks and noise, can be done in a similar manner.

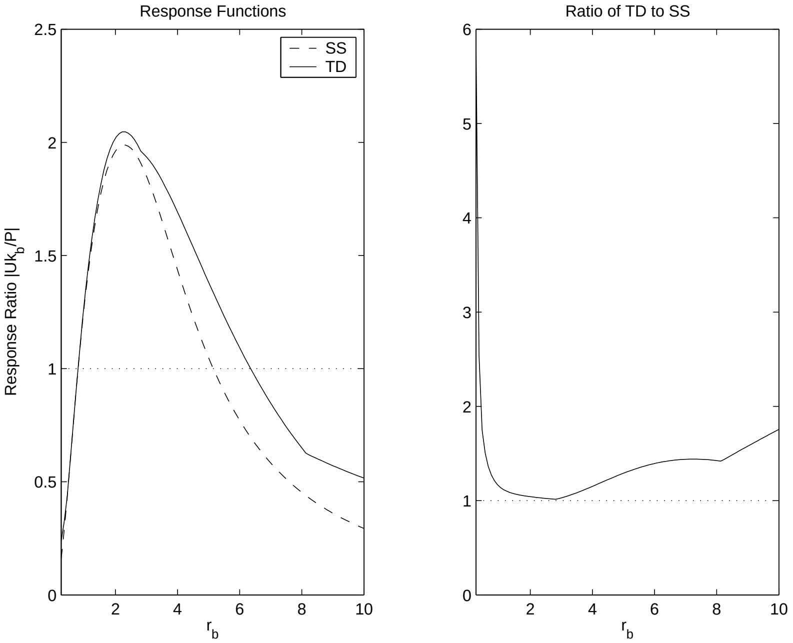

Figure 10.7: The left plot consists of normalized displacement curves for the combined system at an apical site on the BM. The right plot is the ratio of the maximum transient response to the steady state response.

In practice, however, the time domain middle ear has significantly reduced the dispersive instability for inputs including clicks, noise, and speech.

Chapter 11

The Active Cochlea

11.1 The linear active TM/BM model

To make the linear passive TM/BM model (9.25) active, a self-excited vibrational force acting on the BM is added as follows:

$$M_p^f \xi_{tt} + C_p \xi_t + K_p \xi = F + F_a$$

where

$$F_a = \begin{bmatrix} \gamma[c_4(u_t - v_t) + k_4(u - v)] \\ 0 \end{bmatrix}$$

The difference $u - v$ represents outer hair cell (OHC) displacement. The parameter $\gamma \in [0, 1]$ is the active gain control and at this point is constant. Later, to achieve nonlinear compression along with other nonlinear effects, we will make the active gain a nonlinear nonlocal functional of BM displacement. Bringing F_a to the left, we have

$$M_p^f \xi_{tt} + (C_p - \gamma C_a)\xi_t + (K_p - \gamma K_a)\xi = F \tag{11.1}$$

where

$$C_a = \begin{bmatrix} c_4 & -c_4 \\ 0 & 0 \end{bmatrix}, \quad K_a = \begin{bmatrix} k_4 & -k_4 \\ 0 & 0 \end{bmatrix}$$

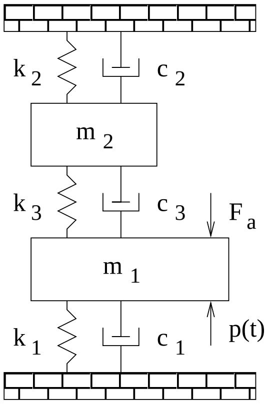

Figure 11.1: This is similar to Figure 9.5, except now we have a self-excited vibrational force F_a acting on the BM (m_1).

11.1.1 Steady State Formulation

Referring to equation (9.21), we see all we need for the steady state model is the impedance function Z_{CP}. This is given by

$$Z_{CP} = Z_p + Z_a$$

where Z_p and Z_a are

$$Z_p = Z_1 + \frac{Z_2 Z_3}{Z_2 + Z_3}$$

$$Z_a = -\gamma \frac{Z_2 Z_4}{Z_2 + Z_3}$$

and $Z_4 = k_4/i\omega + c_4$. See equation (9.24) for definitions of the remaining Z_i. Note that Z_p is Z_{CP} for the passive TM/BM model given in (9.24). For our parameters (see Table A.2), we have $\Re\{Z_a\} < 0$, and thus Z_a gives us *negative damping.* It is effectively adding energy to the system. Combining the passive and active impedances together, we arrive at the impedance for the linear active system as

$$Z_{CP} = Z_1 + \frac{Z_2(Z_3 - \gamma Z_4)}{Z_2 + Z_3}$$

11.1.2 Dispersion Relation

Following the same analysis as in Section 9.3.2 for the linear passive case, we have the following (considering all damping constants equal to zero):

$$\begin{aligned}
\omega_1(0) &= 0 \\
\omega_2(0) &= \pm\sqrt{\frac{k_2+k_3}{m_2}} \\
\omega_1'(0) &= \pm\sqrt{\frac{H\left(k_1+\frac{k_2(k_3-\gamma k_4)}{k_2+k_3}\right)}{2\rho}} \\
\omega_2'(\kappa) &= \kappa\left(\frac{H}{2\rho}\right)\sqrt{\frac{m_2}{k_2+k_3}}\left[\frac{k_3(k_3-\gamma k_4)}{k_2+k_3}\right]+O(\kappa^2)
\end{aligned}$$

11.2 Nonlinear Nonlocal Active Gain

We will now finally move from the linear to the nonlinear regime. The nonlinearity will exist in the active gain and will depend on the absolute BM displacement. The idea is that if the BM displacement is “small”, then the cochlea should have full active gain and amplify the response ($\gamma = 1$). Analogously, if the BM displacement is “large”, then the cochlea should have no active gain ($\gamma = 0$) and essentially behave like the dead, passive cochlea. Thus, a naive approach would be to use the function

$$\gamma(u) = \frac{1}{1+u^2} \tag{11.2}$$

where u is the BM response. The difficulty with this is that in practice, since the BM response has large slopes at various places, the active gain becomes too “discontinuous”, the result being an unstable BM response in time. To correct this, we take a Guassian average of the BM response centered at each point and compute (11.2) of this averaged quantity. Formally, we have

$$\hat{u}(x,t) = \frac{2}{\sqrt{\lambda\pi}} \int_0^L e^{-(x-s)^2/\lambda} u^2(s,t)\ ds$$

and gain

$$\gamma(u,x,t) = \frac{1}{1+\theta\hat{u}}$$

where θ, λ are constants. This makes the nonlinearity nonlocal in nature. The assumption here is that the feedback process at each point on the BM knows about the full BM response in a weighted fashion [8].

11.3 Numerics

11.3.1 Semi-Discrete Formulation

We will discretize in space first to arrive at a semi-discrete formulation of the model equations (11.1), giving

$$M\vec{\xi}_{tt} + C(t)\vec{\xi}_t + K(t)\vec{\xi} = \vec{b}(t)$$

where

$$M = \begin{bmatrix} M_1 + \alpha M_f & 0 \\ 0 & M_2 \end{bmatrix}, \quad \alpha = \frac{4\rho}{N-1}$$

$$C(t) = C_p - \hat{\Gamma}(t)C_a = \begin{bmatrix} C_1 + C_3 - \Gamma(t)C_4 & -(C_3 - \Gamma(t)C_4) \\ -C_3 & C_2 + C_3 \end{bmatrix}$$

$$K(t) = K_p - \hat{\Gamma}(t)K_a = \begin{bmatrix} K_1 + K_3 - \Gamma(t)K_4 & -(K_3 - \Gamma(t)K_4) \\ -K_3 & K_2 + K_3 \end{bmatrix}$$

$$\vec{b}(t) = \begin{bmatrix} 2\rho s_{tt}(t)(\vec{x} - L) \\ 0 \end{bmatrix}$$

$$M_{f,ij} = \sum_{k=1}^{K} Q_2(\beta_k)\cos(\beta_k x_i)\cos(\beta_k x_j) w_j$$

C_p, K_p, C_a, and K_a are now block diagonal, where $K_i = \text{diag}\{k_i\}$ and $C_i = \text{diag}\{c_i\}$. Also, $M_i = \text{diag}\{m_i\}$, $\Gamma(t) = \text{diag}\{\gamma_i(t)\}$ and $\hat{\Gamma}(t) = \text{diag}\{\Gamma(t), 0\}$. The numbers w_j are numerical integration weights in the discretization of the integral and are chosen based on the desired degree of accuracy. Note that we can write $M_f = M_f^s W$, where $W = \text{diag}(w_j)$ and M_f^s is symmetric and positive definite. The matrix M_f is essentially the fluid load on the BM and *dynamically couples* the system.

11.3.2 Iterative Scheme

Using a second order approximation of the first and second derivatives, we arrive at

$$(L_p - L_a^n)\vec{\xi}^{n+1} = \vec{B}^n \Longrightarrow \vec{\xi}^{n+1,k+1} = L_p^{-1}\vec{B}^n + L_p^{-1}L_a^n\vec{\xi}^{n+1,k} \tag{11.3}$$

where superscript n denotes discrete time, k denotes iteration and

$$L_p = 2M + \frac{3}{2}\Delta t C_p + \Delta t^2 K_p$$

$$L_a^n = \hat{\Gamma}^n[\frac{3}{2}\Delta t C_a + \Delta t^2 K_a]$$

$$\vec{B}^n = \Delta t^2 \vec{b}(n\Delta t) + M(5\vec{\xi}^n - 4\vec{\xi}^{n-1} + \vec{\xi}^{n-2}) + \frac{\Delta t}{2}C^n(4\vec{\xi}^n - \vec{\xi}^{n-1}).$$

11.3.3 System Reduction

The previous iterative scheme is a $2N \times 2N$ system. Using the block matrix structure of the system, we can reduce the iteration to an $N \times N$ system. Along these lines, we write L_p and L_a^n in block matrix form as

$$L_p = \begin{bmatrix} \tilde{M}_1 & -P_3 \\ -P_3 & \tilde{M}_2 \end{bmatrix}$$

$$L_a^n = \left[\begin{array}{cc} \Gamma^n P_4 & -\Gamma^n P_4 \\ 0 & 0 \end{array} \right]$$

where

$$\begin{aligned} \tilde{M}_1 &= 2(\alpha M_f + M_1) + P_1 + P_3 \\ \tilde{M}_2 &= 2M_2 + P_2 + P_3 \\ P_i &= \frac{3}{2}\Delta t C_i + \Delta t^2 K_i \end{aligned}$$

It is easily seen that the left inverse of L_p is given by

$$L_p^{-1} = \left[\begin{array}{cc} D^{-1} & D^{-1}\tilde{M}_2^{-1}P_3 \\ \tilde{M}_2^{-1}P_3 D^{-1} & \tilde{M}_2^{-1}P_3 D^{-1}\tilde{M}_1 P_3^{-1} \end{array} \right]$$

where

$$\begin{aligned} D &= \tilde{M}_1 - P_3\tilde{M}_2^{-1}P_3 \\ &= 2\alpha M_f + [2M_1 + P_1 + P_3(I - \tilde{M}_2^{-1}P_3)] \\ &= \{2\alpha M_f^s + [2M_1 + P_1 + P_3(I - \tilde{M}_2^{-1}P_3)]W^{-1}\}W \\ &\equiv D_s W \end{aligned}$$

Note that D is invertible since M_f^s is positive definite, thus invertible, and all other terms are positive diagonal matrices, and thus their sum is positive definite and invertible. We also have

$$L_p^{-1}L_a^n = \left[\begin{array}{cc} D^{-1}\Gamma^n P_4 & -D^{-1}\Gamma^n P_4 \\ \tilde{M}_2^{-1}P_3 D^{-1}\Gamma^n P_4 & -\tilde{M}_2^{-1}P_3 D^{-1}\Gamma^n P_4 \end{array} \right]$$

Letting $\vec{B}^n = (\vec{B}_1^n, \vec{B}_2^n)$, we have

$$\begin{aligned} \vec{u}^{n+1,k+1} &= W^{-1}[\zeta_1^n + D_s^{-1}\Gamma^n P_4(\vec{u} - \vec{v})^{n+1,k}] \\ \vec{v}^{n+1,k+1} &= W^{-1}\tilde{M}_2^{-1}P_3[\zeta_2^n + D_s^{-1}\Gamma^n P_4(\vec{u} - \vec{v})^{n+1,k}] \end{aligned}$$

where

$$\zeta_1^n = D_s^{-1}[\vec{B}_1^n + \tilde{M}_2^{-1}P_3\vec{B}_2^n]$$
$$\zeta_2^n = D_s^{-1}[\vec{B}_1^n + \tilde{M}_1 P_3^{-1}\vec{B}_2^n]$$

11.3.4 Convergence of Iterative Scheme

We need the following Lemma:

Lemma 11.3.1. *If*

$$M = \begin{bmatrix} A & -A \\ B & -B \end{bmatrix}$$

where M is $2N \times 2N$ and A, B are $N \times N$, then every non-zero eigenvalue of M is an eigenvalue of $A - B$.

Proof:

Let λ be a non-zero eigenvalue of M with non-trivial eigenvector $\vec{x} = (\vec{x}_1, \vec{x}_2)$, where $\vec{x}_i$ are $N \times 1$. Thus, $M\vec{x} = \lambda\vec{x}$ gives

$$A(\vec{x}_1 - \vec{x}_2) = \lambda\vec{x}_1 \tag{11.4}$$

$$B(\vec{x}_1 - \vec{x}_2) = \lambda\vec{x}_2 \tag{11.5}$$

Subtracting the two equations, we have

$$(A - B)(\vec{x}_1 - \vec{x}_2) = \lambda(\vec{x}_1 - \vec{x}_2)$$

Now, if $\vec{x}_1 - \vec{x}_2 = 0$, then from above and $\lambda \neq 0$, we have $\vec{x}_1 = \vec{x}_2 = 0$. But this means $\vec{x} = 0$, which is a contradiction. Thus, λ is an eigenvalue of $A - B$ with non-trivial eigenvector $\vec{x}_1 - \vec{x}_2$. $\square$

Theorem 11.3.2. *There exists a constant $C > 0$ such that if $\Delta t < C$, then*

$$\rho(L_p^{-1} L_a^n) < 1$$

where ρ is the spectral radius. Thus, the iterative scheme (11.3) converges.

Proof:

By the above lemma applied to above, with constant γ, we have

$$\begin{aligned}\sigma(L_p^{-1}L_a^n) &\subset \gamma\sigma(D^{-1}P_4 - \tilde{M}_2^{-1}P_3D^{-1}P_4) \\ &= \gamma\sigma[(I - \tilde{M}_2^{-1}P_3)D^{-1}P_4]\end{aligned}$$

where σ denotes spectrum. Thus, we have

$$\begin{aligned}\rho(L_p^{-1}L_a^n) &\leq \gamma||(I - \tilde{M}_2^{-1}P_3)W^{-1}D_s^{-1}P_4||_2 \\ &\leq \gamma||(I - \tilde{M}_2^{-1}P_3)W^{-1}||_2||D_s^{-1}||_2||P_4||_2\end{aligned}$$

Now, let $(\lambda, \vec{x})$ be the eigen-pair of D_s with λ the smallest eigenvalue and $||\vec{x}|| = 1$. Note that $\lambda > 0$ since D_s is positive definite. Thus, we have $1/\lambda$ is the largest eigenvalue of D_s^{-1}, which gives

$$||D_s^{-1}||_2 \leq 1/\lambda$$

Thus, using the definition of D_s, we have

$$\begin{aligned}\lambda &= \vec{x}^T D_s \vec{x} \\ &= \vec{x}^T\{2\alpha M_f^s + [2M_1 + P_1 + P_3(I - \tilde{M}_2^{-1}P_3)]W^{-1}\}\vec{x} \\ &\geq \vec{x}^T\{[2M_1 + P_1 + P_3(I - \tilde{M}_2^{-1}P_3)]W^{-1}\}\vec{x} \\ &\geq \min\{[2m_1 + p_1 + p_3(1 - \tilde{m}_2^{-1}p_3)]w^{-1}\}\end{aligned}$$

where lowercase represents diagonal entries. The third line above follows from $2\alpha M_f^s$ being positive definite. Finally, we have

$$\begin{aligned}\rho(L_p^{-1}L_a^n) &\leq \gamma||(I - \tilde{M}_2^{-1}P_3)W^{-1}||_2||D_s^{-1}||_2||P_4||_2 \\ &\leq \gamma\frac{\max[(1 - \tilde{m}_2^{-1}p_3)w^{-1}]\max(p_4)}{\min\{[2m_1 + p_1 + p_3(1 - \tilde{m}_2^{-1}p_3)]w^{-1}\}}\end{aligned}$$

For Δt small enough, we have convergence. □

With our parameters, for convergence it is *sufficient* that $\Delta t \leq 0.0008$. In practice, however, convergence is seen for Δt as large as 0.01.

Chapter 12

Nonlinearities & Results

12.1 Isointensity Curves

The dynamic range and nonlinear compressive effects of the ear are demonstrated with our model using iso-intensity curves. We will compare our results to the iso-intensity curves disussed and displayed in section 8.4.5.1. Figure 12.1 shows iso-intensity curves at $x = 0.77$ cm with characteristic frequency (CF) 10 kHz. The left plot is the linear steady state active case. The parameter is the active gain γ, and for each value of the active gain we get a curve that is a function of the input frequency. It is basically an output/input ratio and gives the transfer characteristics of the ear at that particular active level. Notice that when $\gamma = 1$, the BM at the characteristic place is most sensitive at the corresponding characteristic frequency, but at lower values of the gain, the sensitivity peak shifts to lower frequencies.

Analogously, the second plot in Figure 12.1 shows isointensity curves for the nonlinear active time domain model where now the parameter is the intensity of the input stimulus in dB SPL (sound pressure level). For the time domain, we measure the root-mean-square BM amplitude from 5 ms (to remove transients) up to a certain time T. Note that for high-intensity tones, the model becomes passive while low-intensity tones give a more active model. This shows *compression.* Again, there is a frequency shift of the sensitivity peak (about one-half octave) from low to high-intensity stimuli in agreement

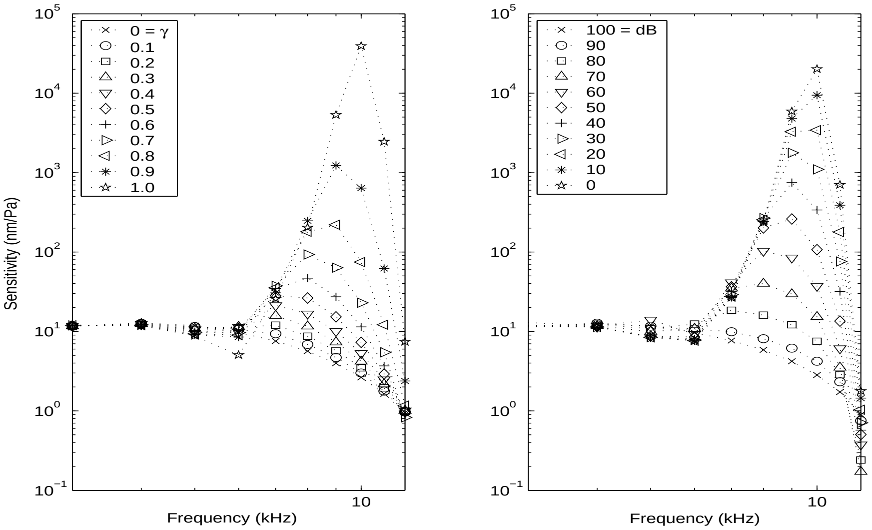

Figure 12.1: Both figures are sensitivity curves for CP = 0.77 cm or CF = 10 kHz. The left plot is a collection of sensitivity curves for the linear steady state active model where the parameter is the active gain γ. The right plot is a collection of sensitivity curves for the nonlinear time domain model where the parameter is pressure at the eardrum in dB SPL (sound pressure level).

with [29], so called half-octave shift. The plot agrees well with Figure 8.5.

12.2 Complex Stimuli

The first non-sinusoidal input we look at is a click. In the experiment in the left plot of Figure 12.2, we put probes at varying characteristic places associated with frequencies ranging from 0.5-4 kHz to measure the time series BM displacement. The click was 40 dB with duration 0.1 ms starting at 0.4 ms. All responses were normalized to amplitude 1. The plot is similar to Figure 4 in [10]. In the right plot of Figure 12.2, a probe was placed at CP for 6.4 kHz and the time series BM volume velocity was recorded for various intensities and the sensitivity plotted. This shows, similar to Figure 12.1, the

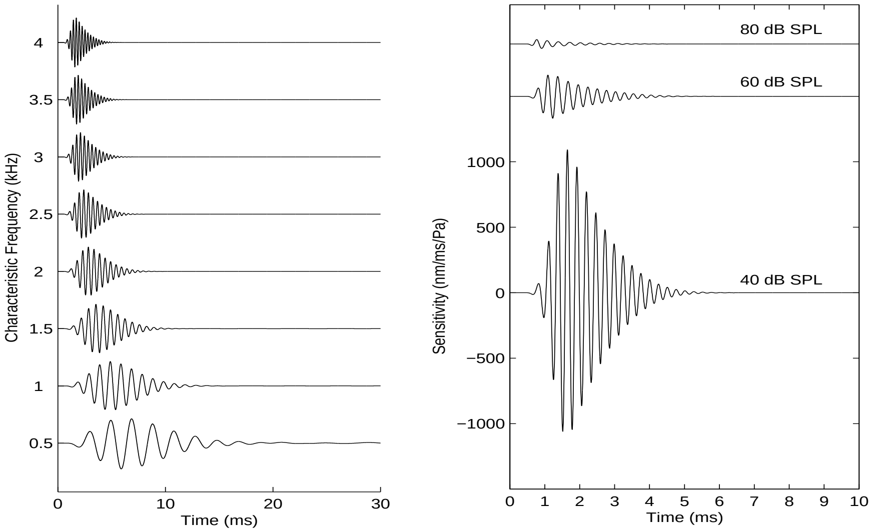

Figure 12.2: An impulse, or click, lasting 0.1 ms starting at 0.4 ms is input into the nonlinear nonlocal ear model. The left plot is the BM displacement time series for various CF's ranging from 0.5-4 kHz. The right plot is a sensitivity plot for various stimulus intensities at CF = 6.4 kHz.

compression effects at higher intensities. This is very similar to Figure 8.6.

The second non-sinusoidal input we explore is Gaussian white noise. Figure 12.3 is similar in all regards to Figure 12.2. Notice again in the right plot the compression effect.

12.3 Combination Difference Tones

As was mentioned in section 8.4.5.2, any nonlinear system with multiple sinusoidal inputs will create difference tones. If two frequencies f_1 and f_2 are put into the ear, $nf_1 \pm mf_2$ will be created at varying intensities, where n and m are nonnegative integers. The cubic difference tone, denoted $f = 2f_1 - f_2$, where $f_1 < f_2$, is the most prominent. Figure 12.4 contains three plots of one

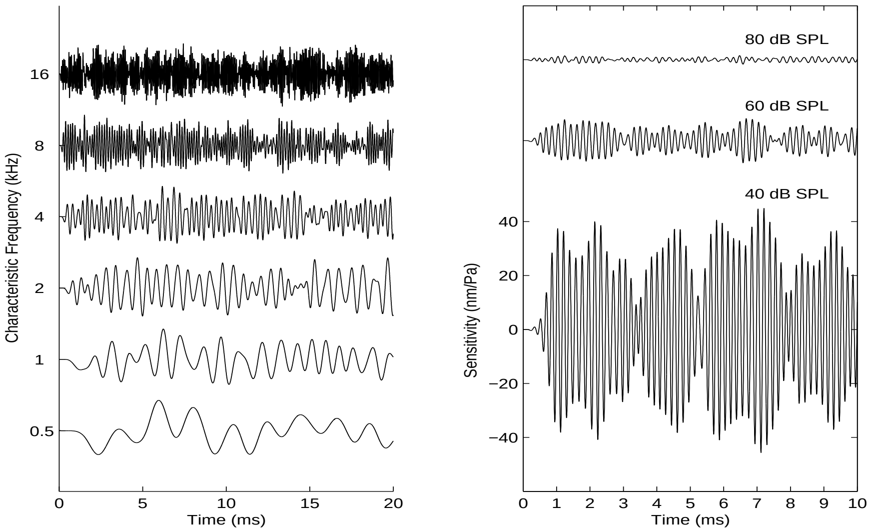

Figure 12.3: Gaussian noise is input into the ear. The left plot is the BM displacement times series for various CF's ranging from 0.5-16 kHz. The right plot is a sensitivity plot for CF = 6.4 kHz.

experiment. The experiment consists of two sinusoidal tones, 7 and 10 kHz at 80 dB each. The cubic difference tone is 4 kHz. The plot on the left is the BM profile for the experiment at 15 ms. We see combination tone peaks at 1.21 cm (CP for 4 kHz), 1.54 cm (CP for 2 kHz) and 1.85 cm (CP for 1 kHz). The middle plot shows the snapshot at 15 ms of the active gain parameter, showing the difference tones getting an active boost. Finally, the right plot is a spectrum plot of the time series for BM displacement at 1.21 cm, the characteristic place for 4 kHz. The cubic difference tone is above 1 nm and can therefore be heard. This is comparable to Figure 8.7. In Figure 8.7, there are a lot more difference tones since the two frequencies f_1 and f_2 are very close to one another.

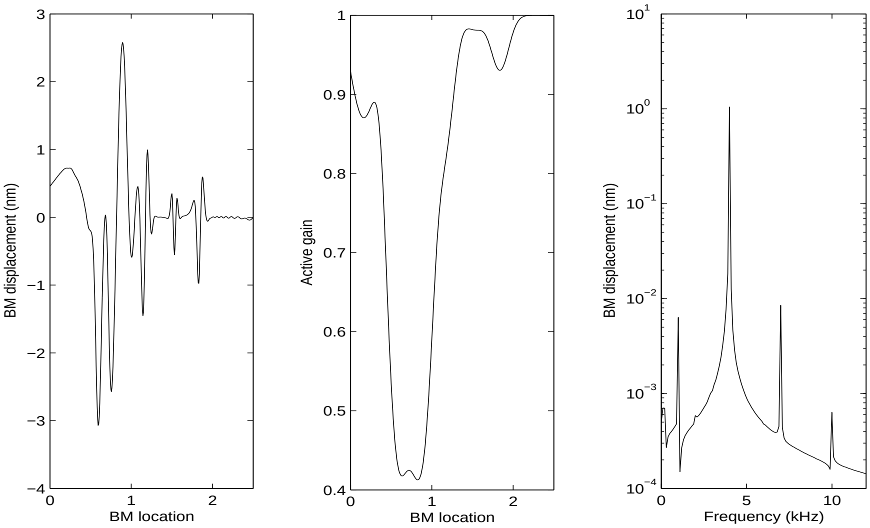

Figure 12.4: Two sinusoidal tones, 7 and 10 kHz at 80 dB each, are the input. The left and middle plots are snapshots at 15 ms of BM displacement and active gain, respectively. The right plot is a spectrum plot of the BM displacement time series at CP for 4 kHz.

12.4 Multi-tone Suppression

Two-tone (and multi-tone) suppression is characteristic of a compressive nonlinearity and has been recognized in the ear [9, 12, 29]. Figure 12.5 illustrates two-tone suppression and is a collection of isodisplacement curves that show decreased tuning in the presence of suppressors and is similar to Figure 8.8 reproduced from [29]. We placed a probe at the CP for 4 kHz (1.21 cm) and input sinusoids of various frequencies. At each frequency, we record the pressure at the eardrum that gives a 1 nm displacement for 4 kHz in the FFT spectrum of the time series response at CP. The curve without suppressors is dashed with circles. We then input each frequency again, but this time in the presence of a low side (0.5 kHz) tone and high side (7.5 kHz) tone, both at 80 dB. Notice the reduced tuning near the CF. Also notice the asymmetry

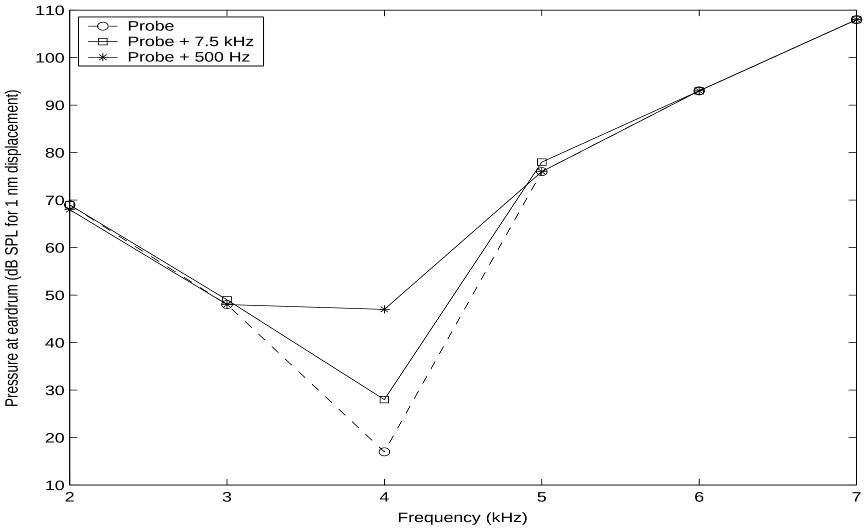

Figure 12.5: Isodisplacement curves at CP for 4 kHz showing responses to single tones (dashed line w/ circle) and responses to the same tones in the presence of high-side and low-side suppressors presented at 80 dB SPL.

of suppression, which shows low side is more suppressive than high side, in agreement with [12].

For multi-tone suppression, we look at tonal suppression of noise. In Figure 12.6, for each plot, a probe was placed at every grid point along the BM and the time response was measured from 15 ms up to 25 ms. The signal in each consisted of noise at 50 dB with a 2 kHz tone ranging from 40 dB to 80 dB (top to bottom). An FFT was performed for each response and its characteristic frequency FFT amplitude was recorded and plotted in decibels relative to the average of the response spectrum of 0 dB noise from 0.5-16 kHz. We see suppression of all frequencies, with again low-side suppression stronger than high-side suppression. Figure 12.6 is qualitatively similar to Figure 3 in [9]. It is useful to compare this figure with Figure 12.7. This figure is the same as Figure 12.6, except we do an FFT of the input signal at the eardrum.

Comparing these two figures shows that we have a new spectral transform that can be used in place of an FFT in certain applications, for example signal recognition and noise suppression.

12.5 Cochleograms

We also tested our model on the speech waveform "Mat" taken from the word "Matlab". We constructed a cochleogram as follows (similar to a spectrogram):

1. Find the frequency range $F = [f_0, f_1]$ of the sound based on its sampling frequency.
2. Construct the characteristic place-to-frequency map $\mathcal{W}(x)$ using iso-input curves (see equation (8.3)). Place probes at all points $x = \mathcal{W}^{-1}(F)$ on the BM. This gives linear-scale in BM location and log-scale in frequency (also called Bark scale).
3. We then start with a linear-scale in frequency and interpolate the inverse map $\mathcal{W}^{-1}$ to create the characteristic frequency-to-place map $\mathcal{X}$ with linear-scale in frequency and log-scale in BM location. Round each log-scale BM location to the nearest linear-scale BM location. This gives a map $\mathcal{X}$ that goes from linear-scale frequency to linear-scale BM location.
4. Start the time-domain simulation. At varying times based on the window size and amount of window overlap, compute the FFT of the windowed time-domain waveform BM responses at each probe. For each frequency f from (3), pick out the frequency f from the FFT at its CP. waveform.

This will give the response FFT for a time-windowed slice. This is one column of the spectrogram.

5. Repeat (4).

Figure 12.8 shows one spectrogram followed by three cochleograms. Each cochleogram has a 13.5 ms window size with a 12.8 ms window overlap. The first plot on the left is a regular FFT spectrogram of the input waveform. We can see the consonants ‘m’ and ‘t’ at the beginning and end of the spectrogram. The vowel ‘a’ is in the middle and is most pronounced. The remaining three are cochleograms for varying input SPL. Midrange to loud SPLs create more nonlinear effects, and thus in the 60 dB SPL plot we see extra microstructures in the cochleogram. Notice how similar (as least to the eye) the loud SPL cochleogram is to the FFT spectrogram, showing how relatively well the FFT does in processing speech in the passive case. It is conjectured that the nonlinear effects in the midrange intensities will help in speech recognition in the presence of noise (see Figure 12.6). This will be explored in more detail in future work.

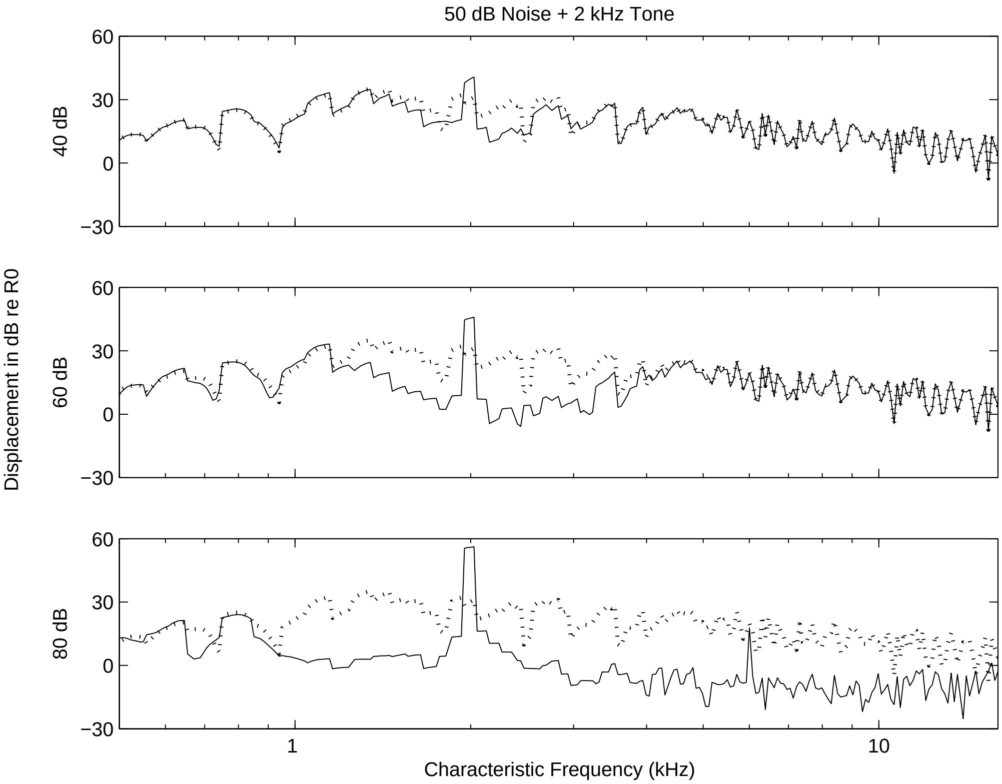

Figure 12.6: Spectrum plots of BM responses for characteristic frequencies along the BM, from 500 Hz to 16 kHz, with 50 dB noise and a 2 kHz tone ranging from 40-80 dB. R0 is the average of the BM response spectrum of 0 dB noise from 0.5-16 kHz. The solid line represents noise with tone, the dotted line noise without tone.

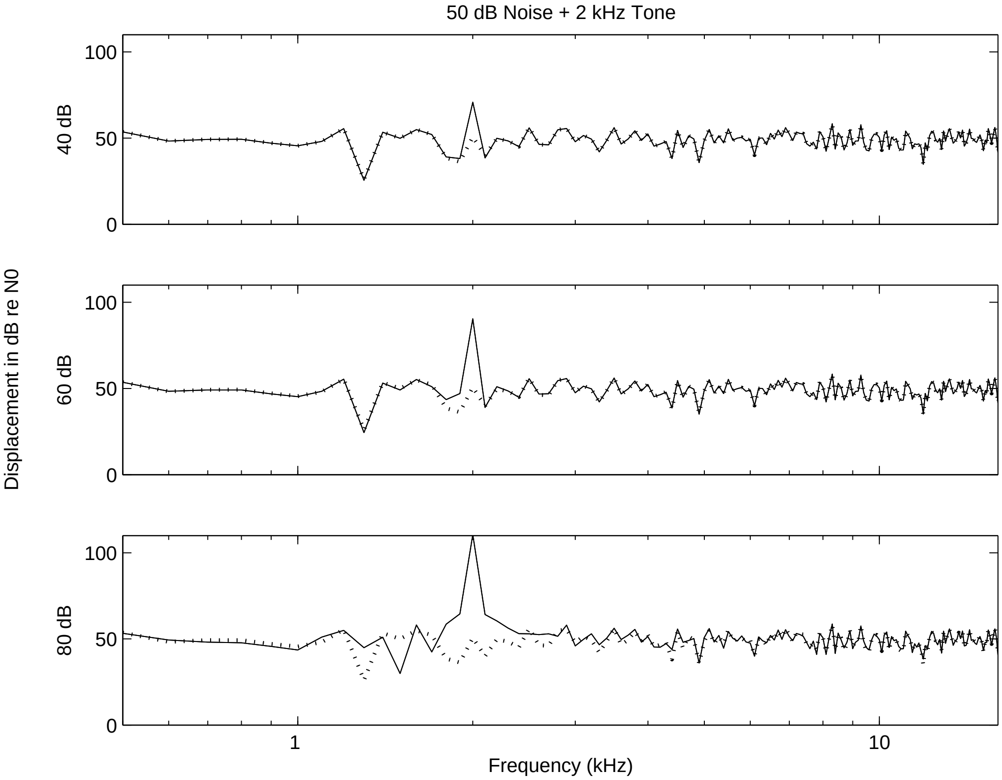

Figure 12.7: Spectrum plots of input signals consisting of 50 dB noise and a 2 kHz tone ranging from 40-80 dB. N0 is the average of the spectrum of 0 dB noise from 0.5-16 kHz. The solid line represents noise with tone, the dotted line noise without tone.

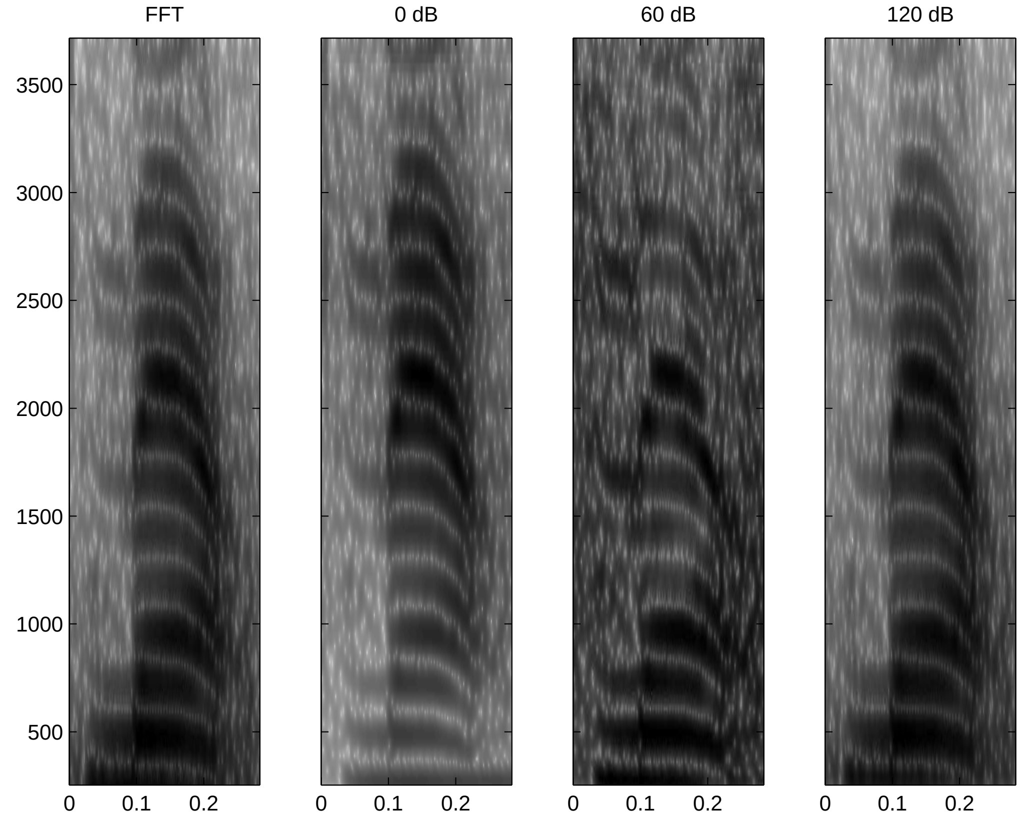

Figure 12.8: The left plot is the FFT spectrogram of the speech waveform 'Mat' from 'Matlab'. The other three plots are the BM response cochleograms at various SPL. (Black is high energy; white is low energy)

Chapter 13

Conclusion

In conclusion, we have shown that moving from the frequency domain cochlear models to the time domain creates a dispersive instability even using the simplest of models for the cochlear partition. We showed that the time domain middle ear model removes the dispersive instability, showing that the middle ear acts as a transient absorber. We highlighted the differences between a passive and active cochlea, and started with the linear active model of Neely and Kim [25]. We modified their active gain to be a nonlinear nonlocal functional of basilar membrane displacement, and then developed and implemented an efficient and accurate numerical method. Our simulation results were shown to recover many important nonlinear effects in hearing.

There are many avenues for future work. One interesting avenue is to study the inverse problem [31] of finding efficient and automated ways to tune the model to different physiological data, as well as applying the model to psychoacoustic signal processing. More understanding of the higher auditory processes of perception and their coupling with the cochlear mechanics is of particular interest. The use of cochleograms in noise suppression is another avenue. Finally, simplifications on the numerics will be explored that will bring the numerical method a little closer to the time scale for hearing. Currently, one-half second of speech takes 30 minutes of computational time on a 1.7 GHz personal laptop, so there is much room for improvement.

Appendices

Appendix A

Parameters and Tuning

A.1 Vocal Folds

Table A.1: Vocal fold model parameters in cgs units

m_1	0.17 g
m_2	0.03 g
$x_2 - x_1$	0.2 cm
$x_1 - x_0$	0.05 cm
$k_{1,\text{open}}$	45 kdynes
$k_{1,\text{closed}}$	180 kdynes
$y_{0,1}$	0 cm
$k_{2,\text{open}}$	8 kdynes
$k_{2,\text{closed}}$	32 kdynes
$y_{0,2}$	0.0 cm
k_{12}	25 kdynes
$y_{0,12}$	0 cm
y_c	0.001 cm
$A(x_0, t)$	2 cm
$r_{1,\text{open}}$	17.5 dynes/(cm s)
$r_{1,\text{closed}}$	192.4 dynes/(cm s)
$r_{2,\text{open}}$	18.6 dynes/(cm s)
$r_{2,\text{closed}}$	49.6 dynes/(cm s)

A.2 Cochlea

A.2.1 Model Parameters

We start with a modification of the parameters in Neely and Kim [25] (See Table A.2). It is known that higher dimensional models give higher

Table A.2: Cochlear model parameters in cgs units

$m_1(x)$	$3 \cdot 10^{-3}$ g $\cdot$ cm^{-2}	m_e	$6.7 \cdot 10^{-3}$ g $\cdot$ cm^{-2}
$c_1(x)$	$20 + 1500\mathrm{e}^{-2x}$ dyn $\cdot$ s $\cdot$ cm^{-3}	c_e	$2.36 \cdot 10^{2}$ dyn $\cdot$ s $\cdot$ cm^{-3}
$k_1(x)$	$1.1 \cdot 10^{9}\mathrm{e}^{-4x}$ dyn $\cdot$ cm^{-3}	k_e	$4.23 \cdot 10^{6}$ dyn $\cdot$ cm^{-3}
$m_2(x)$	$0.5 \cdot 10^{-3}$ g $\cdot$ cm^{-2}	L	2.5 cm
$c_2(x)$	$10\mathrm{e}^{-2.2x}$ dyn $\cdot$ s $\cdot$ cm^{-3}	H	0.1 cm
$k_2(x)$	$7 \cdot 10^{6}\mathrm{e}^{-4.4x}$ dyn $\cdot$ cm^{-3}	ρ	0.1 g $\cdot$ cm^{-2}
$c_3(x)$	$2\mathrm{e}^{-0.8x}$ dyn $\cdot$ s $\cdot$ cm^{-3}	θ	0.5
$k_3(x)$	$10^{7}\mathrm{e}^{-4x}$ dyn $\cdot$ cm^{-3}	λ	0.08 cm
$c_4(x)$	$1040\mathrm{e}^{-2x}$ dyn $\cdot$ s $\cdot$ cm^{-3}	Δt	$2.5 \cdot 10^{-6} - 10^{-5}$ s
$k_4(x)$	$6.15 \cdot 10^{8}\mathrm{e}^{-4x}$ dyn $\cdot$ cm^{-3}	N	401

sensitivity. This is the case with this model. The 1-D model [25] gives a 90 dB active gain at 16 kHz, whereas the 2-D model gives a 160 dB active gain. Active gain is defined as in [29] as the ratio (in dB) of the maximum steady state response in the fully active case to the maximum steady state response in the fully passive case. We need to tune the system in order to reduce the gain. There are many ways to do this, and the method we choose is to increase all the damping coefficients in Table A.2 by the following:

$$2\mathrm{e}^{0.2773x} c_i \mapsto c_i, \quad i = 1, 2, 3, 4$$

A.2.2 Normalized Displacement for TM/BM

A.2.2.1 Passive Cochlea

Our goal is to derive the normalized displacement curves for the TM/BM system without the fluid and use them to tune the cochlea. We need to be careful, however, in ignoring the fluid since it will change the tuning curves. Continuing anyways, consider the TM/BM passive model (9.23). Since the tuning curves are local BM plots, we fix an x along the BM. We will be substituting harmonic inputs of the form

$$u(x,t) = Ue^{i\omega t}, \quad v(x,t) = Ve^{i\omega t}, \quad p(t) = Pe^{i\omega t}$$

For ease of reading, we will give the nondimensionalized parameters before we begin. They are given as:

$$\omega_B = \sqrt{\frac{k_1}{m_1}}, \quad \omega_T = \sqrt{\frac{k_2}{m_2}}, \quad \omega_{OB} = \sqrt{\frac{k_3}{m_1}}, \quad \omega_{OT} = \sqrt{\frac{k_3}{m_2}}$$

$$r = \frac{\omega}{\omega_B}, \quad \tau = \frac{\omega_T}{\omega_B}, \quad o_B = \frac{\omega_{OB}}{\omega_B}, \quad o_T = \frac{\omega_{OT}}{\omega_B}$$

$$d_{ij} = \frac{c_i}{m_j}, \quad \zeta_{ij} = \frac{c_i}{2m_j\omega_B}$$

Thus, we have

$$(-\omega^2 M_p + i\omega C_p + K_p)\xi = F$$

which gives

$$\begin{bmatrix} P \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} (k_1 + k_3 - m_1\omega^2) + i\omega(c_1 + c_3) & -k_3 - i\omega c_3 \\ -k_3 - i\omega c_3 & (k_2 + k_3 - m_2\omega^2) + i\omega(c_2 + c_3) \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix}$$

Dividing the first equation by m_1 and the second by m_2, we get

$$\begin{bmatrix} P/m_1 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} (\omega_B^2 + \omega_{OB}^2 - \omega^2) + i\omega(d_{11} + d_{31}) & -\omega_{OB}^2 - i\omega d_{31} \\ -\omega_{OT}^2 - i\omega d_{32} & (\omega_T^2 + \omega_{OT}^2 - \omega^2) + i\omega(d_{22} + d_{32}) \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix}$$

Dividing both equations by ω_B, we have

$$\begin{bmatrix} P/k_1 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} (1 + o_B^2 - r^2) + 2ir(\zeta_{11} + \zeta_{31}) & -o_B^2 - 2ir\zeta_{31} \\ -o_T^2 - 2ir\zeta_{32} & (\tau^2 + o_T^2 - r^2) + 2ir(\zeta_{22} + \zeta_{32}) \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix}$$

giving

$$\begin{bmatrix} U \\ V \end{bmatrix} =$$

$$\frac{\begin{bmatrix} (\tau^2 + o_T^2 - r^2) + 2ir(\zeta_{22} + \zeta_{32}) & o_B^2 + 2ir\zeta_{31} \\ o_T^2 + 2ir\zeta_{32} & (1 + o_B^2 - r^2) + 2ir(\zeta_{11} + \zeta_{31}) \end{bmatrix} \begin{bmatrix} P/k_1 \\ 0 \end{bmatrix}}{\det(\Lambda)}$$

where Λ is the nondimensionalized matrix. We are interested in the expression

$$\frac{Uk_1}{P}.$$

The numerator of the above ratio is given by

$$\tau^2 + o_T^2 - r^2 + 2ir(\zeta_{22} + \zeta_{32})$$

The denominator is given by $\det(\Lambda) = R_p + iI_p$, where

$$\begin{aligned} R_p &= \mathrm{Re}(\det(\Lambda)) \\ &= (1 + o_B^2 - r^2)(\tau^2 + o_T^2 - r^2) - o_B^2 o_T^2 - 4r^2[\zeta_{11}\zeta_{22} + (\zeta_{11} + \frac{o_B^2}{o_T^2}\zeta_{22})\zeta_{32}] \end{aligned}$$

$$\begin{aligned} I_p &= \mathrm{Im}(\det(\Lambda)) \\ &= 2r[(1 + o_B^2 - r^2)(\zeta_{22} + \zeta_{32}) + (\tau^2 + o_T^2 - r^2)(\zeta_{11} + \frac{o_B^2}{o_T^2}\zeta_{32}) - 2o_B^2\zeta_{32}] \end{aligned}$$

We eliminated the parameter ζ_{31} using the relation

$$\zeta_{31} = \frac{o_B^2}{o_T^2}\zeta_{32} = \frac{m_2}{m_1}\zeta_{32}$$

Finally, we have the normalized displacement for the TM/BM model

$$\left|\frac{Uk_1}{P}\right| = \sqrt{\frac{(\tau^2 + o_T^2 - r^2)^2 + [2r(\zeta_{22} + \zeta_{32})]^2}{R_p^2 + I_p^2}} \tag{A.1}$$

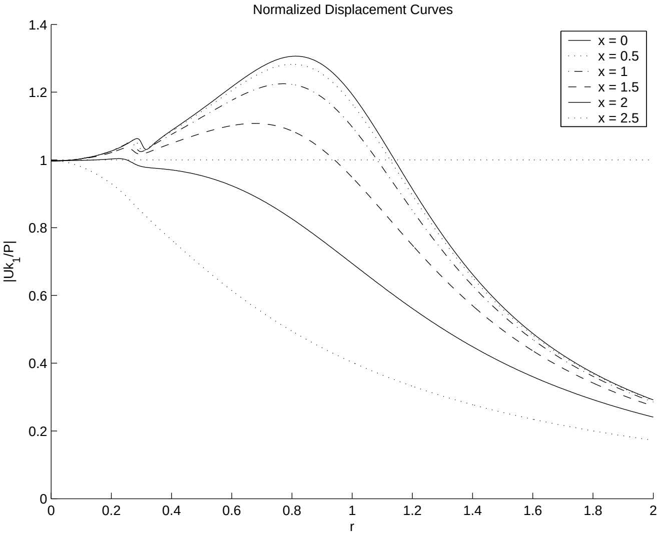

Figure A.1: Plots of normalized displacement curves for the passive TM/BM model at various locations along the BM.

See Figure A.1 for plots of normalized displacement curves over various locations along the BM. Notice that they are very similar to the simple BM normalized displacement curves shown in Figure 10.4. The difference is in the interval $r \in [0.26, 0.3]$ where the second filter's (TM) resonance point is located. The TM resonant point is given from the equations above by

$$r = \sqrt{\tau^2 + o_T^2}$$

Note that, using the parameters from Table A.2, o_T and o_B are constants.

For completeness, we can compute the OHC and TM responses as follows

$$\left|\frac{(U-V)k_1}{P}\right| = \sqrt{\frac{(\tau^2 - r^2)^2 + (2r\zeta_{22})^2}{R_p^2 + I_p^2}}$$

$$\left|\frac{Vk_1}{P}\right| = \sqrt{\frac{o_T^4 + (2r\zeta_{32})^2}{R_p^2 + I_p^2}}$$

A.2.2.2 Active Cochlea

We now consider the TM/BM active model given by equations (11.1). We add the following terms:

$$\omega_A = \sqrt{\frac{k_4}{m_1}},\ \eta = \frac{\omega_A}{\omega_B},\ \zeta_{41} = \frac{c_4}{2m_1\omega_B}$$

A slight modification of the calculation above for the passive case gives

$$\left|\frac{Uk_1}{P}\right| = \sqrt{\frac{(\tau^2 + o_T^2 - r^2)^2 + [2r(\zeta_{22} + \zeta_{32})]^2}{R_a^2 + I_a^2}}$$

where

$$\begin{aligned} R_a &= (1 + o_B^2 - \gamma\eta^2 - r^2)(\tau^2 + o_T^2 - r^2) - (o_B^2 - \gamma\eta^2)o_T^2 - \\ &\quad 4r^2[\zeta_{22}(\zeta_{11} - \gamma\zeta_{41}) + (\zeta_{11} + \frac{o_B^2}{o_T^2}\zeta_{22})\zeta_{32}] \\ I_a &= 2r[(1 + o_B^2 - \gamma\eta^2 - r^2)(\zeta_{22} + \zeta_{32}) + \\ &\quad (\tau^2 + o_T^2 - r^2)(\zeta_{11} + \frac{o_B^2}{o_T^2}\zeta_{32} - \gamma\zeta_{41}) - 2o_B^2\zeta_{32} + \gamma(\eta^2\zeta_{32} + o_T^2\zeta_{41})] \end{aligned}$$

See figure A.2 for plots of the normalized displacement curves with $\gamma = 1$ over various locations along the BM. Again, $\gamma = 0$ gives us equation (A.1). Note that using the parameters in Table A.2, η and ζ_{41} are constants. This is useful in terms of modifying the active tuning of the system.

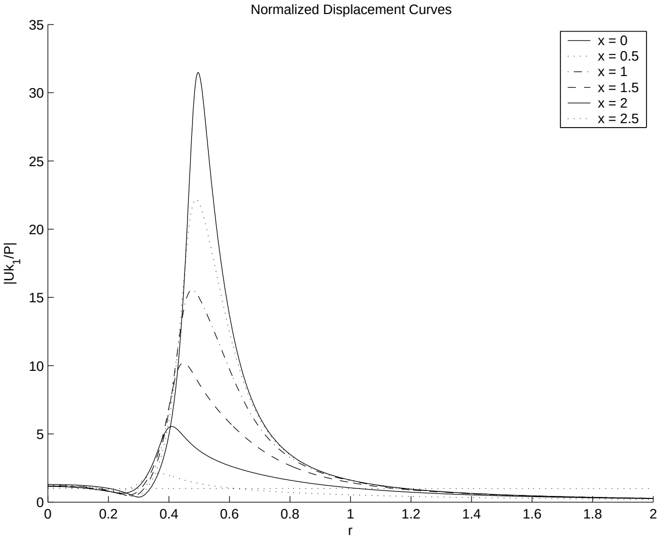

Figure A.2: Plots of normalized displacement curves for the active TM/BM model with $\gamma = 1$ for various locations along the BM.

Bibliography

[1] F. Alipour and R. Scherer. Vocal fold bulging effects on phonation using a biophysical computer model. *J. Voice*, 14:470–483, 2000.

[2] J. B. Allen. Nonlinear cochlear signal processing. In Anthony F. Jahn and Joseph Santos-Sacchi, editors, *Physiology of the Ear*, chapter 19, pages 393–442. Singular Thomson Learning, second edition, 2001.

[3] S. Austin and I. Titze. The effect of subglottal resonance upon vocal fold vibration. *J. Voice*, 11:391–402, 1997.

[4] Gary J. Baker. *Pressure-Feedforward and Piezoelectric Amplification Models for the Cochlea.* PhD thesis, Stanford University, March 2000.

[5] I. Bogaert. Speech production by means of a hydrodynamic model and a discrete-time description. Technical Report 1000, Inst Perception Res, Einhoven, the Netherlands, 1994.

[6] Peter Dallos, Arthur N. Popper, and Richard R. Fay, editors. *The Cochlea.* Springer-Verlag, 1996.

[7] E. de Boer. Mechanics of the cochlea: Modeling efforts. In A. Popper P. Dallos and R. Fay, editors, *The Cochlea*, pages 258–317. Springer Handbook of Auditory Research, Springer-Verlag, 1996.

[8] E. de Boer and A. L. Nuttall. Properties of amplifying elements in the cochlea. In A. W. Gummer, editor, *Biophysics of the Cochlea: From Molecules to Models.* Proc. Internat. Symp., Titisee, Germany, 2002.

[9] L. Deng and C. D. Geisler. Responses of auditory-nerve fibers to multiple-tone complexes. *J. Acoust. Soc. Amer.*, 82(6):1989–2000, 1987.

[10] L. Deng and Issam Kheirallah. Numerical property and efficient solution of a transmission-line model for basilar membrane wave motions. *Signal Processing*, 33(3):269–285, September 1993.

[11] J. Flanagan. *Speech Analysis, Synthesis and Perception.* Springer-Verlag, New York, Berlin, second edition, 1972.

[12] C. D. Geisler. *From Sound to Synapse.* Oxford University Press, Oxford, 1998.

[13] J. J. Guinan and W. T. Peake. Middle-ear characteristics of anesthetized cats. *J. Acoust. Soc. Amer.*, 41(5):1237–1261, 1967.

[14] A. W. Gummer and B. M. Johnstone. State of stress within the basilar membrane: a re-evaluation of the membrane misnomer. *Hear Res*, 12:353–366, 1983.

[15] H. Helmholtz. *On the Sensations of Tone.* Dover Publications, New York, 1863.

[16] Daniel J. Inman. *Engineering Vibration.* Prentice Hall, second edition, 2000.

[17] K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *AT&T Tech. J.*, 51:1233–1268, 1972.

[18] J. Jiang and I. Titze. Measurement of vocal fold intraglottal pressure and impact stress. *J. Voice*, 8:132–144, 1994.

[19] M. D. LaMar, Y-Y. Qi, and J. Xin. Modeling vocal fold motion with a hydrodynamic semi-continuum model. *J. Acoust. Soc. Amer.*, 114:455–464, 2003.

[20] M. D. LaMar, Y-Y. Qi, and J. Xin. Signal processing of acoustic signals in the time domain with an active nonlinear nonlocal cochlear model. *arXiv: q-bio.QM/0411032*, 2004 (Submitted to Signal Processing).

[21] R. LeVeque. *Numerical Methods for Conservation Laws.* Birkhauser-Verlag, Basel, 1990.

[22] R. LeVeque. *Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm*, volume 146. 1998.

[23] L. Mongeau, N. Franchek, C. Coker, and R. Kubil. Characteristics of a pulsating jet through a small modulated oriface, with applications to voice production. *J. Acoust. Soc. Amer.*, 102:1121–1132, 1997.

[24] Stephen T. Neely. Mathematical modeling of cochlear mechanics. *J. Acoust. Soc. Am.*, 78(1):345–352, July 1985.

[25] Stephen T. Neely and D. O. Kim. A model for active elements in cochlear biomechanics. *J. Acoust. Soc. Am.*, 79(5):1472–1480, May 1986.

[26] X. Pelorson, A. Hirschberg, A. Wijnands, and H. Bailliet. Description of the flow through *in vitro* models of the glottis during phonation. *Acta Acust. (Beijing)*, 3:191–202, 1995.

[27] K. Pohlmann. *Principles of Digital Audio.* McGraw-Hill Video/Audio Professional, fourth edition, 2000.

[28] L. Rabiner and B-H. Juang. *Fundamentals of Speech Recognition.* Prentice Hall PTR, 1993.

[29] Luis Robles and Mario A. Ruggero. Mechanics of the mammalian cochlea. *Physiological Reviews*, 81(3):1305–1352, July 2001.

[30] R. Sataloff. The human voice. *Sci. Am.*, 1992:108–115, December 1992.

[31] M. Sondhi. The acoustical inverse problem for the cochlea. In M. Holmes and L. Rubenfeld, editors, *Lecture Notes in Biomathematics*, pages 95–104. Springer-Verlag, 1980.

[32] K. Stevens. *Acoustic Phonetics.* MIT, Cambridge, MA, 2000.

[33] B. Story and I. Titze. Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Amer.*, 97:1249–1260, 1995.

[34] I. Titze. Current topics in voice production mechanisms. *Acta Oto-Laryngol*, 113:421–427, 1993.

[35] L. Voldrich. Mechanical properties of the basilar membrane. *Acta Oto-laryngol*, 86:331–335, 1978.

[36] G. von Békésy. *Experiments in Hearing.* McGraw-Hill, New York, 1960.

[37] G. B. Whitham. *Linear and Nonlinear Waves.* Wiley, New York, 1974.

[38] J. Xin. Dispersive instability and its minimization in time domain computation of steady state responses of cochlear models. *J. Acoust. Soc. Amer.*, 115(5, Pt. 1):2173–2177, 2004.

[39] J. Xin, Y-Y. Qi, and L. Deng. Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing. *Communication in Math Sciences*, 1(2):211–227, 2003.

Vita

Michael Drew LaMar was born in Fairfax, Virginia on August 13, 1975, to John Michael LaMar and Ruth Early LaMar. He began undergraduate studies at University of Texas at San Antonio in fall of 1993, where he started doing undergraduate research with Dr. Mary Lou Zeeman. He obtained his bachelors of science degree in spring of 1997. He attended one year of graduate school at University of California at Los Angeles before transferring to the University of Texas at Austin in the fall of 1998. He began studies with his advisor, Dr. Jack Xin, in the spring of 2001.

Permanent address: 12303A Cahone Trail
Austin, Texas 78729

This Dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.